

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: December 1, 2004, 22:49:23 ; Search time 107 Seconds

(without alignments)

2238.653 Million cell updates/sec

Title: US-09-867-570-2

Perfect score: 1763 1 MESKSSWVIRLGLSLMDST.....EGGGLPQETLELSGRLEQ 337

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xip
-Q=/cgn2_1/uspro.spool.p/us09867570/runat.01122004.161123.6373/app.query.fasta_1.519
-DB=Issued Patents NA -QPM=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MTN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09867570 @CGN1.1.105 @runat.01122004.161123.6373 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPELCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Issued Patents NA.*

1:	/cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2:	/cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3:	/cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4:	/cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5:	/cgn2_6/ptodata/1/ina/PCRTUS.COMB.seq.*
6:	/cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1642	93.1	969	4	US-09-254-227A-4
2	1598	90.6	969	4	US-09-254-227A-6
3	1405	79.7	969	4	US-09-254-227A-8
4	1382	78.4	969	4	US-09-254-227A-10
5	1382	78.4	969	4	US-09-254-227A-12
6	1360	77.1	969	4	US-09-254-227A-14
7	834.5	47.3	1011	4	US-09-254-227A-2
8	529	30.0	291	4	US-09-495-050A-164
9	496	28.1	275	4	US-09-016-434-330
10	461	26.1	2416	4	US-09-016-434-1264
11	430.5	24.4	1388	4	US-09-016-434-1225
12	430.5	24.4	1388	5	PCT-US93-06251-26

13	392.5	22.3	1327	6	US-09-464-970B-134	Patent No. 5320941
14	381.5	21.6	2435	4	US-08-961-825-5	Sequence 134, App
15	255.5	14.5	2911	3	US-09-480-784-5	Sequence 5, Appl1
16	255.5	14.5	2911	3	US-09-480-784-5	Sequence 5, Appl1
17	240	13.6	1200	5	PCT-US95-03032-1	Sequence 1, Appl1
18	238	13.5	1062	4	US-09-016-434-1207	Sequence 1, Appl1
19	231	13.1	1200	5	PCT-US92-02977-1	Sequence 1, Appl1
20	227.5	12.9	1106	5	PCT-US92-02977-5	Sequence 5, Appl1
21	227.5	12.9	1106	5	PCT-US95-03032-4	Sequence 4, Appl1
22	226.5	12.8	1510	4	US-07-759-568-4	Sequence 1140, Ap
23	226.5	12.8	1510	4	US-09-023-655-1140	Sequence 8, Appl1
24	226.5	12.8	1750	4	US-08-023-655-1135	Sequence 1135, Ap
25	226.5	12.8	1143	1	US-08-467-125-1	Sequence 1, Appl1
26	225.5	12.8	1143	3	US-08-911-320A-1	Sequence 1, Appl1
27	225.5	12.8	1143	3	US-09-217-101-1	Sequence 1, Appl1
28	225.5	12.8	1143	3	US-09-016-434-46	Sequence 46, Appl
29	225.5	12.8	1842	4	US-09-944-807-1	Sequence 7, Appl1
30	220	12.5	1244	1	US-08-417-103-7	Sequence 7, Appl1
31	220	12.5	1244	1	PCT-US92-02977-6	Sequence 6, Appl1
32	220	12.5	1373	5	PCT-US95-03032-3	Sequence 3, Appl1
33	220	12.5	1373	5	US-09-016-434-1206	Sequence 1, Appl1
34	217	12.3	1058	4	US-09-944-807-1	Sequence 1083, Ap
35	217	12.3	1910	4	US-09-023-655-1083	Sequence 1446, Ap
36	217	12.3	1317	4	US-09-016-434-1446	Sequence 126, Ap
37	214.5	12.2	1438	4	US-09-576-1608-12	Sequence 12, Appl
38	214	12.1	1062	4	US-09-170-496D-1	Sequence 1, Appl1
39	211	12.0	1062	4	US-09-170-496D-1	Sequence 163, App
40	210	11.9	1068	4	US-09-745-842-1	Sequence 1, Appl1
41	206	11.7	1068	4	US-09-170-496D-7	Sequence 7, Appl1
42	205.5	11.7	1289	4	US-09-016-434-1430	Sequence 1430, Ap
43	205	11.6	1002	4	US-09-170-496D-167	Sequence 167, App
44	205	11.6	1265	4		
45	204	11.6	1002	4		

ALIGNMENTS

★ RESULT 1
US-09-254-227A-4
Sequence 4, Application US/09254227A
Patent No. 6696257
GENERAL INFORMATION:
APPLICANT: Ahmad, Sultan
APPLICANT: Banville, Denis
APPLICANT: Fortin, Yves
APPLICANT: Lembo, Paola
APPLICANT: O'Donnell, Paola
APPLICANT: Shi-Hsiang, Shen
TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
FILE REFERENCE: 81823/268117
CURRENT APPLICATION NUMBER: US/09/254,227A
CURRENT FILING DATE: 1999-03-03
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin version 3.0
SEQ ID NO 4
LENGTH: 969
TYPE: DNA
ORGANISM: Homo sapiens
US-09-254-227A-4

Alignment Scores:
Pred. No.: 1.07e-157
Score: 1642.00
Percent Similarity: 98.76%
Best local Similarity: 97.20%
Query Match: 93.14%
DB: 4
Gaps: 0

US-09-867-570-2 (1-337) x US-09-254-227A-4 (1-969)

Cy 16 MetAspSerThrIleProValIleuGlyThrGluLeuThrProIleAsnGlyIleGluGlu 35
Db 1 ATGATCAACACATCCAGTCTTGAGTACAAACCTGACACCATCAACGAGCGAGAG 60

```

QY 36 ThrProCysTyrLeuGlnThrLeuSerPheThrGlyLeuThrCysAlaValSerLeuVal 55
Db 61 ACTCCTGTGCTACAAACCAAAACCCCTGACCTTCAAGGGGCTGACGGACATCACTCCCTGTGC 120
QY 56 AlaLeuThrGlyAsnAlaValLeuThrPheLeuGlyCysArgMetArgArgAsnAla 75
Db 121 GCGCTGACAGGAAACGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95
Db 181 GCTCCATCTACATCTCAACCTGAGTGGCGGCCCACTTCCTCTTACGGCGCCCACTT 240
QY 96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
Db 241 ATATTGTGGCGCTTACCCCTCAATCAATATTCGCCATTCCTCAAAATCTCAAGTCTC 300
QY 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
Db 301 GTGATGACCTTCCCTACTTATAGGCTTAAGGCTTAAGGCTTAAGGCTTAAGGCTTAAGGCT 360
QY 136 CysLeuSerIleLeuThrProIleTyrPheHisCysArgArgProArgTyrLeuSerSer 155
Db 361 TGCCGTGCTCATCTGAGGCGCCATCTGTGATCACTGCGCGCGCCAGATACCTGTATCG 420
QY 156 ValMetCysValLeuLeuThrPheAlaLeuSerLeuLeuArgSerIleLeuGluThrPhePhe 175
Db 421 GTCATGTGTGCTGTGCTGTGCGGCTGTGCTGTGCTGTGCGGAGTATCCCTGAGGAGTGC 480
QY 176 CysAspPheLeuPheSerGlyAlaAspSerValTTPCysGluThrSerAspPheIleThr 195
Db 481 TGTGATCTCTGTTAGTGTGCTAATCTGTTTGTGAGGAAAGTCAGATTTCAATTACA 540
QY 196 IleAlaThrPheValPheLeuCysValValLeuCysGlySerSerLeuValLeuLeuVal 215
Db 541 ATCCGGTGGCTGGTGTGTTTATGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 600
QY 216 ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValThrIleLeuLeu 235
Db 601 AGGATCTCTGTGATCCCGGAAAGATGCCGCTTACCAAGGCTTACGATCACTCTCTCTC 660
QY 236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnThrAlaLeuPheSer 255
Db 661 ACAGTGTGCTGTCTCTCTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 720
QY 256 ArgIleHisLeuAspTyrPheValLeuPheCysHisValHisLeuValSerIlePheLeu 275
Db 721 AGGATCCACCTGATGAAAGCTTATTTGTGATGATGATGATGATGATGATGATGATGATG 780
QY 276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGln 295
Db 781 TCCGCTCTTAAACAGACGATCCCAACCCCACTTACTCTTCCGAGGCTCTCTTAAAGCAG 840
QY 296 ArgGlnAsnArgGlnAsnLeuPheValLeuGlnArgAlaLeuGlnAspThrProGln 315
Db 841 CGTCAAAATAGGCAAAACCTGAAGCTGTTCTTCAAAAGGCTCTGACAGGACAGCGCTGAG 900
QY 316 ValAspGluGlyGlyGlyTyrPheLeuProGlnGluThrLeuGluLeuSerGlySerArgLeu 335
Db 901 GTGATGAGAGGTGAGGAGGCTCTCTCAAGAAACCTGAGAGCTGTGCGGAAACCAATTG 960
QY 336 GluGln 337
Db 961 GAGCAG 966

```

RESULT 2
 US-09-254-227A-6
 : Sequence 6, Application US/09254227A
 : Patent No. 6696257

: GENERAL INFORMATION:
 : APPLICANT: Ahmad, Sulcan
 : APPLICANT: Banville, Denis
 : APPLICANT: Fortin, Yves

```

: APPLICANT: Lembo, Paola
: APPLICANT: O'Donnell, Dajan
: APPLICANT: Shi-Hsiang, Shen
: TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
: FILE REFERENCE: 81823/268117
: CURRENT APPLICATION NUMBER: US/09/254,227A
: CURRENT FILING DATE: 1999-03-03
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 6
: LENGTH: 969
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-254-227A-6

Alignment Scores:
Pred. No.: 5,5e-153 Length: 969
Score: 1598.00 Matches: 307
Percent Similarity: 97.204 Conservative: 5
Best Local Similarity: 95.648 Mismatches: 9
Query Match: 90.648 Indels: 0
DB: 4 Gaps: 0

US-09-867-570-2 (1-337) x US-09-254-227A-6 (1-969)
QY 16 MetAspSerThrIlePheProValLeuGlyThrGluLeuThrProIleAsnGlyArgGluGlu 35
Db 1 AGGATCCCAACCGTCCCTCACTTGTGGTACAGAACTGACCAACATCAAGAGAGAGAG 60
QY 36 ThrProCysTyrLeuGlnThrLeuSerPheThrGlyLeuThrCysAlaValSerLeuVal 55
Db 61 ACTCCTGTGCTACAAACCAAAACCCCTGACCTTCAAGGGGCTGACGGACATCACTCCCTGTGC 120
QY 56 AlaLeuThrGlyAsnAlaValLeuThrPheLeuGlyCysArgMetArgArgAsnAla 75
Db 121 GCGCTGACAGGAAACGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95
Db 181 GCTCCATCTACATCTCAACCTGAGTGGCGGCCCACTTCCTCTTACGGCGCCCACTT 240
QY 96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
Db 241 ATATTGTGGCGCTTACCCCTCAATCAATATTCGCCATTCCTCAAAATCTCAAGTCTC 300
QY 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
Db 301 GTGATGACCTTCCCTACTTATAGGCTTAAGGCTTAAGGCTTAAGGCTTAAGGCTTAAGGCT 360
QY 136 CysLeuSerIleLeuThrProIleTyrPheHisCysArgArgProArgTyrLeuSerSer 155
Db 361 TGCCGTGCTCATCTGAGGCGCCATCTGTGATCACTGCGCGCGCCAGATACCTGTATCG 420
QY 156 ValMetCysValLeuLeuThrPheAlaLeuSerLeuLeuArgSerIleLeuGluThrPhePhe 175
Db 421 GTCATGTGTGCTGTGCTGTGCGGCTGTGCTGTGCTGTGCGGAGTATCCCTGAGGAGTGC 480
QY 176 CysAspPheLeuPheSerGlyAlaAspSerValTTPCysGluThrSerAspPheIleThr 195
Db 481 TGTGATCTCTGTTAGTGTGCTAATCTGTTTGTGAGGAAAGTCAGATTTCAATTACA 540
QY 196 IleAlaThrPheValPheLeuCysValValLeuCysGlySerSerLeuValLeuLeuVal 215
Db 541 ATCCGGTGGCTGGTGTGTTTATGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 600
QY 216 ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValThrIleLeuLeu 235
Db 601 AGGATCTCTGTGATCCCGGAAAGATGCCGCTTACCAAGGCTTACGATCACTCTCTC 660
QY 236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnThrAlaLeuPheSer 255
Db 661 ACAGTGTGCTGTCTCTCTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 720

```



```

Db      541  GTGCGGCGCTGATTTTATATGATGCTCTCTGCTTTTCCAGCCCTGCTGCTGCTC 600
Qy      216  Arg11eUeuCysglYserArgLysMetProleuThrxArgLeuTyVal1Thrl1eUeu 235
Db      601  AGAATCTCTGTGGATCCCGAAGATCCGCTGACGAGCTGATATGACCATCTCTCTC 660
Qy      236  ThVal1eUeuVal1PheUeuLeuCyglYLeuProPheGly11egIntPalaUeuPheSer 255
Db      661  ACAGTGCTGCTCTTCTCTCTCTGCGGCTCTCTCTCGGCAATTCGAGGAGGCTTAATTAC 720
Qy      256  Arg11eUeuAspTrpLysVal1eUeuPheCyHisVal1eUeuVal1Ser11eUeuPhe 275
Db      721  AGAATGCACTGATTTTGAAGTCTTATATGTCATATTTATCTGAGTTGATCTCTCTG 780
Qy      276  SerAlaUeuAsnSerSerAlaAsnPro11e1eTyPhePheVal1GlySerPheArg1n 295
Db      781  TCCTCTTAACAAGTATGCGCAACCCCATCATTTACTTCTTCTGAGGCTCTTTAGGCAG 840
Qy      296  ArgGlnAsnArgGlnAsnLeuLysLeuVal1eUeuGlnArgAlaUeuGlnAspThrProGlu 315
Db      841  CGTCAAAATAGCGAAGAACTGAAGCTGTTCTCCAGAGGCTCTGACAGCAAGCTGAG 900
Qy      316  ValAspGlnGlyGlyGlyTyTrpLeuProGlnGluThrxLeuLeuSerGlySerArgLeu 335
Db      901  GTGATTAAGTGAAAGGCGACCTTCTCGAGAAAGCTGAGCTGTGCGAAGGAGATTG 960

```

RESULT 6

```

US-09-254-227A-14
; Sequence 14, Application US/09254227A
; Patent No. 6696257
; GENERAL INFORMATION:
; APPLICANT: Ahmad, Sultan
; APPLICANT: Banville, Denis
; APPLICANT: Fortin, Yves
; APPLICANT: Lembo, Paola
; APPLICANT: O'Donnell, Dajan
; APPLICANT: Shi-Hsiang, Shen
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
; FILE REFERENCE: 81823/268117
; CURRENT APPLICATION NUMBER: US/09/254,227A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-254-227A-14

```

Alignment Scores:

```

Pred. No.:      8,11e-129      Length:      969
Score:          1360.00      Matches:      264
Percent Similarity: 89.06%      Conservative: 21
Best Local Similarity: 82.50%      Mismatches: 35
Query Match:    77.14%      Indels:      0
DB:              4          Gaps:          0

```

US-09-867-570-2 (1-337) x US-09-254-227A-14 (1-969)

```

Qy      16  MetAspSerThr1LeProVal1LeuGlyThxGluLeuThrxPro11eAsnGlyArgGluGlu 35
Db      1  ATGATTCACACCGTCCAGCTTTCGGTACAAACTGACCAACATCAACGACGAGAGAG 60
Qy      36  ThrProCysTyTrpYsgIntHrxLeuSerPheThxGlyLeuThrxCys11eVal1SerLeuVal 55
Db      61  ACTCTCTTGCATCAATGACACCTCGAGCTTCAACGGTCTGACGTCATCATTTCTCTTGC 120
Qy      56  AlaLeuThrxGlyAsnAlaVal1AlaLeuTrpLeuGluGlyCysArgMetArgAsnAla 75
Db      121  GGAAGTACAGAGAAACGGGTTGTGCTCTGCTCTGAGGCTACCGCATGCCAGGAAGCT 180
Qy      76  ValSer11eTyTrpLeuAsnLeuVal1AlaAlaAspPheLeuPheLeuSerGlyHis11e 95

```

```

Db      181  GTCTCAATTCATCTTCAACCTGAGCGGACGAGACTTCTCTTCCAGCTTCCAAATT 240
Qy      96  11CysSerProLeuArgLeu11eAsn11eArgHisPro11eSer11eLeuSerPro 115
Db      241  ATAGCTTGGCATTTAGCCCTCATCATATATCAGCATCTCATCCGCAAAATCTCTTCT 300
Qy      116  ValMetThrPheProTyTrpPhe11eGlyLeuSerMetLeuSerAla11eSerThrxGluArg 135
Db      301  GTGATGACCTTTCTCCATCTTATACAGGCTGATGATGCTGAGCCCATCAGACACGAGCGC 360
Qy      136  CysLeuSer11eLeuTrpPro11eTyTrpHisCysArgArgProArgTyTrpLeuSerSer 155
Db      361  TGCCTGCTGTCTGTGGCCCATCTGTATCCGTCGCGCGCCGACACACCTGTCAAGC 420
Qy      156  ValMetCysVal1LeuLeuTrpAlaLeuSerLeuLeuAsnSer11eLeuGluTyTrpMetPhe 175
Db      421  GTGCTGTGTCTCTGCTCTGAGGCTCTCTCTGCTGTATGATGCTGAGTGAAGGTTTC 480
Qy      176  CysAspPheLeuPheSerGlyAlaAspSerVal1TrpCysGluThrxSerAspPhe11eThr 195
Db      481  TGTGACTTCTGTTTATGATGCTGATTTCTAGTTGGTGAACGTCAGATTTCATCCCA 540
Qy      196  11eAlaTrpLeuVal1PheLeuCyVal1Val1eUeuCyglYserSerLeuVal1eUeuVal 215
Db      541  GTGCTGTGGCTGATTTTATATGATGCTCTCTGCTTTCCAGCCCTGCTGCTGCTGCTC 600
Qy      216  Arg11eUeuCysglYserArgLysMetProleuThrxArgLeuTyVal1Thrl1eUeu 235
Db      601  AGAATCTCTGTGGATCCCGAAGATCCGCTGACGAGCTGATGTCGATGACCATCTCTCTC 660
Qy      236  ThVal1eUeuVal1PheUeuLeuCyglYLeuProPheGly11egIntPalaUeuPheSer 255
Db      661  ACAGTGCTGCTCTTCTCTCTCTGCGGCTCTCTCTCGGCAATTCGAGGAGGCTTAATTAC 720
Qy      256  Arg11eUeuAspTrpLysVal1eUeuPheCyHisVal1eUeuVal1Ser11eUeuPhe 275
Db      721  AGAATGCACTGATTTTGAAGTCTTATATGTCATATTTATCTGAGTTGATCTCTCTG 780
Qy      276  SerAlaUeuAsnSerSerAlaAsnPro11e1eTyPhePheVal1GlySerPheArg1n 295
Db      781  TCCTCTTAACAAGTATGCGCAACCCCATCATTTACTTCTTCTGAGGCTCTTTAGGCAG 840
Qy      296  ArgGlnAsnArgGlnAsnLeuLysLeuVal1eUeuGlnArgAlaUeuGlnAspThrProGlu 315
Db      841  CGTCAAAATAGCGAAGAACTGAAGCTGTTCTCCAAAGGCTCTGACAGCAAGCTGAG 900
Qy      316  ValAspGlnGlyGlyGlyTyTrpLeuProGlnGluThrxLeuLeuSerGlySerArgLeu 335
Db      901  GTGATTAAGTGAAAGGCGACCTTCTCGAGAAAGCTGAGCTGTGCGAAGGAGATTG 960

```

RESULT 7

```

US-09-254-227A-2
; Sequence 2, Application US/09254227A
; Patent No. 6696257
; GENERAL INFORMATION:
; APPLICANT: Ahmad, Sultan
; APPLICANT: Banville, Denis
; APPLICANT: Fortin, Yves
; APPLICANT: Lembo, Paola
; APPLICANT: O'Donnell, Dajan
; APPLICANT: Shi-Hsiang, Shen
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
; FILE REFERENCE: 81823/268117
; CURRENT APPLICATION NUMBER: US/09/254,227A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1011
; TYPE: DNA
; ORGANISM: rat
US-09-254-227A-2

```


STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 330:
SEQUENCE CHARACTERISTICS:
LENGTH: 275 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CONNUT01
CLONE: 1909132
US-09-016-434-330

Alignment Scores:
Pred. No.: 7.17e-42 Length: 275
Score: 496.00 Matches: 91
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 28.13% Indels: 0
Gaps: 0

US-09-867-570-2 (1-337) x US-09-016-434-330 (1-275)

QY 106 ArgHisProIleSerLysIleLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
DB 1 CGCCATCCATCTCCAAATCTCTGATCTGTGATGACCTTCCCTTATAGGCTTA 60

QY 126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTyrProIleTyr 145
DB 61 AGCATGCTGAGCGGCATCAGACGAGCGGCTGCTGCATCCGTGGCCATCTGGTAC 120

QY 146 HisCysArgArgProAlaGlyTyrLeuSerSerValMetCysValIleLeuTyrPal 165
DB 121 CACTGCGCGCGCCCAATCACTGTCACTGGTCACTGTGCTGCTGCGCCCTGCTCC 180

QY 166 LeuLeuArgSerIleLeuGluTyrMetPheCysAspPheLeuPheSerGlyAlaAspSer 185
DB 181 CTGCTGCGAGTATCTCGAGTGAATGTTCTGTACTTCTGTTAGTGTGCTGATTTCT 240

QY 186 ValTyrCysGluThrSerAspPheIleThrIle 196
DB 241 GTTGTGTGAAACGTGAGATTTCATTAACATC 273

RESULT 10
US-09-016-434-1264
Sequence 1264, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1264:
SEQUENCE CHARACTERISTICS:
LENGTH: 2416 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g244209
US-09-016-434-1264

Alignment Scores:
Pred. No.: 5.98e-37 Length: 2416
Score: 461.00 Matches: 107
Percent Similarity: 57.80% Conservative: 56
Best Local Similarity: 37.94% Mismatches: 81
Query Match: 26.15% Indels: 38
Gaps: 10

US-09-867-570-2 (1-337) x US-09-016-434-1264 (1-2416)

QY 51 IleValSerLeuValAlaLeuThrGlyAsnAlaValLeuTyrPheLeuGlyCysArg 70
DB 1246 CTGATCTCCCTCTGTGGGGCTTATTGAATGGACATCTCTTGGCTTGTGCTGG 1305

QY 71 MetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeu 90
DB 1306 ---GCCACGAATCCCTCATGATGATACATCTCCACCTGGTGGCTGACGATGATAT 1362

QY 91 LeuSerLysIleIleIleCysSerProLeuArgLeuLeuIleArg----- 106
DB 1363 CTT-----TCTGCTCGGAGTGGGCTTTTACAGGACTTGTCTACTATAT 1410

QY 107 HisProIle-----SerLysIleLeuSerProValMetThrPhe 119
DB 1411 CATGAGTCCGCTTTTATCCGATTCCTGCGCATTTGTCTCC-----TTC 1461

QY 120 ProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIle 139
DB 1462 TCCTTGAGGTGTCTCTGCTCTCTGCTGAGCCATCAGACAGCGGTGTGTGTGTC 1521

QY 140 LeuTyrProIleTyrPheCysArgArgProArgTyrLeuSerSerValMetCysVal 159
DB 140 LeuTyrProIleTyrPheCysArgArgProArgTyrLeuSerSerValMetCysVal 159

```

Db 1522 CTCCTCCCATCTGGATGAGATGCCACCGCCCAATAATACATCTAATGTGTCTGACCC 1581
Qy 160 LeuLeuTTPAlaLeuSerLeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeu 179
Db 1582 CTCATCTGGGGCTGCTTTTTCATCAACATAGTAATAATCACTTTTCTTAAGCTTC--- 1638
Qy 180 PheSerGlyAlaAspSerValTrpCysGluThrSerAsp-----PheIleThrIle 196
Db 1639 -----TGGAAACATGTAAAGCATGTGTCTAATTTCTTAAGCTT 1677
Qy 197 AlaTrpLeu-----ValPheLeuCysValIleLeuCysGlySerSerLeuValLeuLeu 214
Db 1678 TCTGGGCTCTTCATGCTATCTCTTTCATCTGTGATGTGTGTGCGAGTCTGACTTACCTC 1737
Qy 215 ValArgIleLeuCysGlySerArgGlySerMetProLeuThrArgLeuTrpValThrIleLeu 234
Db 1738 ATTAAATTCCTGTGTCTGCTCCACGACGAAAGGCCACCAAGGCTCATGTGGCGTGTGCGAG 1797
Qy 235 LeuThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGluTrpAlaLeuPhe 254
Db 1798 ATCTGGCCCCCATGTCTCTACTCTGTGGCCCTTACCCCTGAGCGTG----- 1842
Qy 255 SerArgIleHisLeuAspTrpIleValLeuPheCysHisValHisLeuValSerIlePhe 274
Db 1843 GCACCCCTCATACAGATTTCATAAATGTGTGTACACACCTCTCATTTAATTTCTTGTTC 1902
Qy 275 LeuSerAlaLeuAsnSerSerAlaAsnProIleIleTrpPhePheValGlySerPheArg 294
Db 1903 CTC---ATTATTAACAGACGACGCCCAACCTATCATTTATTTCTTTTGTGGGACCTCAAG 1959
Qy 295 GluArgGluAsnArgGluAsnLeuValLeuValLeuGluAlaLeuGluAsnArgPro 314
Db 1960 AAGAAAGGCTGACAGGAATCTCTCAGAGTGATTTCTCAAGCGGCTTACGAGTAAGCA 2019
Qy 315 GluVal 316
Db 2020 GAGGTG 2025

RESULT 11
US-09-016-434-1225
; Sequence 1225, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice An-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1225:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1388 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9187388
; US-09-016-434-1225

Alignment Scores:
Pred. No.: 3,35e-34 Length: 1388
Score: 430.50 Matches: 103
Percent Similarity: 58.46% Conservative: 56
Best Local Similarity: 37.87% Mismatches: 94
Query Match: 24,42% Indels: 19
DB: Gaps: 8

US-09-867-570-2 (1-337) x US-09-016-434-1225 (1-1388)
Qy 52 ValSerLeuValAlaLeuThrGluValAlaValIleuTrpLeuGlyCysArgMet 71
Db 391 ATCTCCCAAGTGGGGGTTTGTGAGATGGGATTCCTCTGTGCTTCTGCTTCCGATG 450
Qy 72 ArgArgAsnAlaValSerIleTrpIleLeuAsnLeuValAlaAlaAspPheLeu 91
Db 451 AGAAGAAATCCCTTCACTGTCTACATCACCCACTGTCTATGCGACATCTCACTGCTC 510
Qy 92 SerGlyHisIleGlySer-----ProLeuArgLeuIleAsnIleArg 106
Db 511 TTCTGTATTTTCATCTTGTCTATCGACATGCTTTAGATTAGACTTCTTCTGCCAT 570
Qy 107 HisProIleSerIleLeuSerProValMetThrPheProIleGlyLeuSer 126
Db 571 TACTACACATTTGTCACATTATCAGTACTTTTGTGTGGTCAACAGGCGCTTAT 630
Qy 127 MetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTrpProIleTrpHis 146
Db 631 CTGCTGACCGCCCATTAAGTGTGAGAGTGCCCTGTGACGCTTTTACCCATCTGTACCGA 690
Qy 147 CysArgArgProArgTrpIleSerSerValMetCysValIleLeuTrpAlaLeuSerLeu 166
Db 691 TGCATCGCCCAAGTACAGTCGGCATGTGTGTGCTTCTGTGGCTTCTTCTTGC 750
Qy 167 LeuArgSerIleLeuGluTrpMetPheCys-----AspPheLeuPheSerGly 182
Db 751 TTGGTGACCAACATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 810
Qy 183 AlaAspSerValTrpCysGlu---ThrSerAspPheIleThrIle---AlaTrpLeuVal 200
Db 811 AATGAC-----TGCAGACAGCATCATCTTTATGCGCATCTGCTGTGCTG 861
Qy 201 PheLeuCysValValLeuCysGlySerSerLeuValIleLeuValArgIleLeuCysGly 220
Db 862 TTCACGCCCTCATGCTGTG---TTCACACACATCTTGTGTGTGTGTGTGTGTGTGTGT 918
Qy 221 SerArgIleMetProLeuThrArgLeuTrpValThrIleLeuLeuThrValLeuValPhe 240
Db 919 AGCTGGGCTTCCATCTTCCTCAAGCTTTCATAGATCATGATGATGATGATGATGATGAT 978
Qy 241 LeuLeuCysGlyLeuProPheGlyIleGluTrpAlaLeuPheSerArgIleHisLeuAsp 260
Db 979 CTCATCTGCTATGCGCATGAGACTCTTACCTGCTCTATGAGAT----- 1029
Qy 261 TrpIleValLeuPheCysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSer 280
Db 1030 TGGTCGACC---TTGGGAACCTTACACCATTTCCCTGCTTCTTCCACATCAACACT 1086
Qy 281 SerAlaAsnProIleIleTrpPhePheValGlySerPheArgGluArgGluAsnArgGlu 300

```

Db 1087 AGGCGCAACCTTTCATTCTTCTTGTGGAGACAGTAAAGAGAGATTCAAGAG 1146
Qy 301 AsnLeuLysLeuValLeuGlnArgAlaLeuGlnAsp 312
Db 1147 TCCTTAAAGTTGTTGACCAAGGCTTTCAAAGAT 1182

RESULT 12
PCT-US93-06251-26
Sequence 26, Application PC/TUS9306251
GENERAL INFORMATION:
APPLICANT: Wickstrom, Eric and Rife, Jason P.
TITLE OF INVENTION: Trivial Synthesis of Oligonucleotides Containing
TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: NY
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06251
FILING DATE: 19930630
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8586
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 1388 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-26

Alignment Scores:
Pred. No.: 3,35e-34 Length: 1388
Score: 430.50 Matches: 103
Percent Similarity: 58.46% Conservative: 56
Best Local Similarity: 37.87% Mismatches: 94
Query Match: 24.42% Indels: 19
DB: 5 Gaps: 8

US-09-867-570-2 (1-337) x PCT-US93-06251-26 (1-1388)

Qy 52 ValSerLeuValAlaLeuThrGlyAsnAlaValAlaLeuTrpLeuLeuGlyCysArgMet 71
Db 391 ATCTCCCAAGTGGGGTTTGTGAAGATGGATTCTCTCGGTCCGTCGTCCTCCGATG 450
Qy 72 ArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeu 91
Db 451 AGAAGAAATCCCTTCACTGCTACATCACCACCTGTCTATCGCAGACATCTCACTGCTC 510
Qy 92 SerGlyHisIleIleCysSer-----ProleuArgLeuIleAsnIleArg 106
Db 511 TTCTGATTTTCACTTGTCTATCGACTATGCTTAATATGAGCTTCTTCTGCGCAT 570
Qy 107 HisProIleSerIleLeuSerProValMetThrPheProTyrPheIleGlyLeuSer 126
Db 571 TACTACACAAATGTCACATTATCAGTGACTTTTCTGTGTGGCTACACACAGGGCTCTAT 630

Qy 127 MetLeuSerAlaIleSerThrGlnArgCysLeuSerIleLeuTrpProIleTyrHis 146
Db 631 CTGCTGACGGCCATTAGTGTGAGAGAGTGCCTGTCAGTCTTTTACCCCATCTGTGACGA 690
Qy 147 CysArgArgProArgTyrLeuSerSerValMetCysValLeuLeuTrpAlaLeuSerLeu 166
Db 691 TGCATGCGCCCAAGTACCAAGTCCAGTGGCATTTGGTCTGTCCCTTCTGTGGCTTTTCTTGC 750
Qy 167 LeuArgSerIleLeuGlnTrpMetPheCys-----AspPheLeuPheSerGly 182
Db 751 TTGTGTACACACCATGAGTATGTCATGTCATGCACAGAGAGAGAGAGTCACTTCGCG 810
Qy 183 AlaAspSerValTrpCysGlu---ThrSerAspPheIleThrIle---AlaTrpLeuVal 200
Db 811 AATGAC-----TCCAGACAGTCATCATCTTTTATCCATCCCTGAGGCTTCCTGTC 861
Qy 201 PheLeuCysValAlaLeuCysGlySerSerLeuValLeuValArgIleLeuCysGly 220
Db 862 TTCACGCCCCCTCATGCTGTG---TCCAGACCATCTTGTGTCGTAAGATCCGAGAAC 918
Qy 221 SerArgLysMetProleuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPhe 240
Db 919 ACGTGGGCTTCCCATTCCTCCAGCTTTACATAGTCATGTCACCATCATATATATTC 978
Qy 241 LeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAsp 260
Db 979 CTATCTTCGCTTACCCCATGAGACTCTTATCCTGCTGTGATGATGAT----- 1029
Qy 261 TrpLysValLeuPheCysHisIleLeuValSerIleLeuSerAlaLeuAsnSer 280
Db 1030 TGTGTGACCC---TTTGGAACTTACACCATTTCCCTGCTTCTTCCACATATACACT 1086
Qy 281 SerAlaAsnProIleIleTyrPhePheValGlySerPheArgGlnArgGln 300
Db 1087 AGGCGCAACCTTTCATTACTTCTTGTGGAGACAGTAAAGAGAGATTCACAGAG 1146
Qy 301 AsnLeuLysLeuValLeuGlnArgAlaLeuGlnAsp 312
Db 1147 TCCTTAAAGTTGTTGACCAAGGCTTTCAAAGAT 1182

RESULT 13
5320941-1
Patent No. 5320941
Applicant: Young, Dallan;Wigler, Michael H.;Passano
Otlavio
TITLE OF INVENTION: DNA SEQUENCES ENCODING MMS ONHCOGENE,
POLYPEPTIDES ENCODED THEREFROM AND DIAGNOSTIC AND OTHER METHODS
BASED THEREFROM
NUMBER OF SEQUENCES: 2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/06/872,087
FILING DATE: 06-JUN-1986
SEQ ID NO:1:
LENGTH: 1327
5320941-1

Alignment Scores:
Pred. No.: 2.27e-30 Length: 1327
Score: 392.50 Matches: 97
Percent Similarity: 55.20% Conservative: 57
Best Local Similarity: 34.77% Mismatches: 72
Query Match: 22.26% Indels: 53
DB: 6 Gaps: 9

US-09-867-570-2 (1-337) x 5320941-1 (1-1327)

Qy 52 ValSerLeuValAlaLeuThrGlyAsnAlaValAlaLeuTrpLeuLeuGlyCysArgMet 71
Db 390 ATCTCCCAAGTGGGGTTTGTGAAGATGGATTCTCTCGGTCCGTCGTCCTCCGATG 449
Qy 72 ArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeu 91
Db 450 AGAAGAAATCCCTTCACTGCTACATCACCACCTGTCTATCGCAGACATCTCACTGCTC 509

```

QY 92 SerGlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIys 111
Db 510 -----TTCGTG-----ATTTCACTCTGTCTATGACATATGCTTAAAGATTAT 551
QY 112 IleLeuSer-----ProValMetThr-----PhePro 120
Db 552 GAGCTTTTCTTGCGCATTAACACAAATGTCACATTATCAAGTAGAGTTTCTGTTGGC 611
QY 121 TyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeu 140
Db 612 TACAACACGGGCGCTCATCTGCTGACGCGCATTAAGTGAGAGAGGCTGTCAGACGCTT 671
QY 141 TrpProIleTrpTyrHisCysArgArgProArgTyrLeuSerSerValMetCysValLeu 160
Db 672 TACCCCATCTGTATCCCATGCGCATGCGCCCAAGTACAGTCCAGCTAGGCTGTGGCCCTT 731
QY 161 LeuTrpAlaLeuSerIleLeuArgSerIleLeuGluTrpMetPheCysAspPheLeuPhe 180
Db 732 CTGTGGGCTCTTTCTTGCTGTGAGCACCATGAGATGATGTCATGTCATCCACAGAA 791
QY 181 SerGlyAlaAspSerValTrpCysGluThrSerAsp-----PheIle 194
Db 792 GAAGAGAGTGAAGTCT-----CCGAATGACTGCCACAGCTCATCATCTTTATA 839
QY 195 ThrIle---AlaTrpLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeu 213
Db 840 GCCATCTGAGGCTCTGTGCTTCTGCTGACGCGCTCATCTGCTGCTGCTGCTGCTGCTG 896
QY 214 LeuValArgIleLeuCysGlySerArgIleMetPheLeuThrArgLeuTyrValIle 233
Db 897 GTCCGGAAGATCCCAAGAAACAGTGGGCTTCCCATCTCTCCAAAGCTTTTACATAGCATC 956
QY 234 LeuLeuThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeu 253
Db 957 ATGGTCACCATCATTAATCTCTCATCTTCTGCTATGCGCATGAGACTCTCTTACCTGCTG 1016
QY 254 PheSerArgIleHisLeuAspTrpTyrValLeuPheCysHisValHisLeuValSerIle 273
Db 1017 TAC----- 1019
QY 274 PheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPhe 293
Db 1020 -----TAGAGTAGCGCCAAACCTTTTACTTTCTTTGTGGAGACAGAT 1064
QY 294 ArgGlnArgGlnAsnArgGlnAsnLeuValLeuValLeuGlnArgAlaLeuGlnAsp 312
Db 1065 AAGAAGAGAGATTCAGACAGTCTTAAAGTTGTTCTGACACAGGCGCTTTCAAAGAT 1121

RESULT 14
US-09-484-970B-134
/ Sequence 134, Application US/09484970B
/ Patent No. 6426186
/ GENERAL INFORMATION:
/ APPLICANT: Jones, Karen A.
/ APPLICANT: Volkmut, Wayne
/ APPLICANT: Walker, Michael G.
/ TITLE OF INVENTION: BONE REMODELING GENES
/ FILE REFERENCE: PB-0014 US
/ CURRENT APPLICATION NUMBER: US/09/484, 970B
/ NUMBER OF SEQ ID NOS: 172
/ SOFTWARE: PERL Program
/ SEQ ID NO 134
/ LENGTH: 2435
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: InCyt ID No. 6426186 247789.2CBI
/ LOCATION: 93, 128, 132, 143-144, 2419, 2427, 2429
/ OTHER INFORMATION: a, t, c, g, or other

```

```

US-09-484-970B-134
Alignment Scores:
Pred. No.: 7,15e-29 Length: 2435
Score: 381.50 Matches: 100
Percent Similarity: 52.26% Conservative: 50
Best Local Similarity: 34.84% Mismatches: 112
Query Match: 21.64% Indels: 25
DB: 4 Gaps: 7

US-09-867-570-2 (1-337) x US-09-484-970B-134 (1-2435)
QY 48 LeuThrCysIleValSerIleValAlaLeuThrGlyAsnAlaValLeuTrpLeuLeu 67
Db 642 CTCCTCTGCTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 692
QY 68 GlyCysArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuAlaAlaAsp 87
Db 693 GCGTTCTTCATCAAGAGAACCCCTTCTTCATCTTCTTCATCTTCTTCACCTGCGACGCGCGAT 752
QY 88 PheLeuPheLeuSerGlyHisIleIleCysSerProLeuArg----- 101
Db 753 GTGGGCTACCTTTTACAGACAGGCGGTGTTCTTCATCTGACACAGGAGGCTTCTGAGC 812
QY 813 AGCTTGGCCAGTACATCCGACAGCGGTGCGCGGCTGCGGCTGCGATG----- 863
Db 102 ---LeuIleAsnIleArgHisProIleSerIleLeuSerProValMetThrPhePro 120
QY 121 TyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArgCysLeu-SerIleLe 140
Db 864 TTCCTTACCGCGGTGAGGCTCTGCGCGCGCGGCTCGACCTCGAGCGCTGCGCTGCTGCTAT 923
QY 140 uTrpProIleTrpTyrHisCys-ArgArgProArgTyrLeuSerSerValMetCysVal 160
Db 924 CTTCCTCCGCGGTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 983
QY 160 euLeuTrpAlaLeuSerIleLeuArgSerIleLeuGluTrpMetPheCysAspPheLeuP 180
Db 984 TGCTGTGGGCTGTGCTCTCTCTGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1043
QY 180 heSerGlyAlaAspSerValTrpCysGluThrSerAsp---PheIleThrIleAlaTrp 199
Db 1044 GCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1103
QY 199 euValPheLeuCysValValLeuCysGlySerSerIleValLeuValArgIleLeu 219
Db 1104 TCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1163
QY 219 ySerIleArgIleMetProLeuThr---ArgLeuTyrValIleThrIleLeuLeuVal 238
Db 1164 GCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1223
QY 238 euValPheLeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSerArgIle 258
Db 1224 CCGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1273
QY 258 IseLeuAspTrp-----LysValLeuPheCysHisValHisLeuValSerIlePheLeu 276
Db 1274 -----TGGGCTTCCAGATCCGCGCGCGCGCGCGCGCGCGCGCGAGTACTGACTGCTGCA 1325
QY 276 exAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGln 296
Db 1326 TCTGCATCAACAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1385
QY 296 rGlnAsnArgGlnAsnLeuValLeuValLeuGlnArgAlaLeuGlnAspTrpProGlu 316
Db 1386 AGCGGCTGTGAGCGCGCTCAGGCTGCTCTTCAAGCGGCGCGCTGCGGAGCGGCTGAGC 1445
QY 316 AlaArgGluGlyGly 321
Db 1446 TGGGGAGGCGCGGCGCGC 1462

RESULT 15

```

```

US-08-981-825-5
; Sequence 5, Application US/08981825
; Patent No. 6040426
; GENERAL INFORMATION:
; APPLICANT: OGAWA, KAZUYAKUJI
; APPLICANT: TANAKA, KAZUYA
; APPLICANT: NAGATA, KINYA
; APPLICANT: TAKANO, SYOICHI
; TITLE OF INVENTION: PROTEIN SPECIFIC TO HUMAN
; TITLE OF INVENTION: TN2, GENE (B19) ENCODING THE SAME, AND
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Knobbie, Martens, Olsson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/981,825
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Altman, Daniel E
; REGISTRATION NUMBER: 34,115
; REFERENCE/DOCKET NUMBER: MSMT4_001APC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-760-0404
; TELEFAX: 714-760-9502
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2911 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; US-08-981-825-5

Alignment Scores:
Pred. No.:          5,78e-16      Length:      2911
Score:              255.50        Matches:     83
Percent Similarity: 43.30%       Conservative: 56
Best Local Similarity: 25.86%    Mismatches: 107
Query Match:         14.49%      Indels:      75
BB:                  3           Gaps:         10

US-09-867-570-2 (1-337) x US-08-981-825-5 (1-2911)

QY      51 IlevlserleuvAlalaleuthrglyAenAlvalValleutRpleuLeuglyCyvArg 70
Dy      233 CTGGGCTCGCTGCCTGGCCGTGTGAAGATGAGTCACTCTTCCTGGTGCGGGCGTGGCC 292
Qy      71 MetArGaASnaAlavalserIetyrIleleuAsnleuValAlaalAAPheLey--- 89
Dy      223 ATGCCCGACAGCGGTGTCAACCACACTGGGTGCTGCACTGGCGCGCTGTCCGACTGTGGCC 352
Qy      90 -----PheLeuSer-----Gly 93
Dy      353 TCTGCTCCCTGCCCTCTTTCACTCACTCTTGTGGCGGTGGCCACTGCTGGAGCTGGGC 412
Qy      94 HstllellelcyGsserProleuArGleulleAsnllleArghIsProIIeserylleleu 113
Dy      413 ACCACCTTCTGC-----AAACTGCAC 433

```


This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: December 1, 2004, 22:50:53 ; Search time 553 Seconds

(without alignments)
3348.557 Million cell updates/sec

Title: US-09-867-570-2

Perfect score: 1763
1 MESKSSWYRIGFLSHMDST.....EGGGWLPQETLELSGRLEQ 337

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Egapop 6.0 , Egapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3694831 seqs, 2747406616 residues

Total number of hits satisfying chosen parameters: 7389662

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xip
-O/cgn2 1/USPTO.spool.p/US09867570/runat.01122004.161124.6393/app.query.fasta_1.519
-DB=Published Applications NA -QFMT=fastad -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blonum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=JOCL -OUTFMT=ptc -NORM=ext -HRAPSIZ=500 -MINLEN=0
-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:*
- 5: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:*
- 6: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:*
- 7: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:*
- 8: /cgn2_6/prodata/1/pubpna/US11_NEW_PUB.seq:*
- 9: /cgn2_6/prodata/1/pubpna/US12_NEW_PUB.seq:*
- 10: /cgn2_6/prodata/1/pubpna/US13_NEW_PUB.seq:*
- 11: /cgn2_6/prodata/1/pubpna/US14_NEW_PUB.seq:*
- 12: /cgn2_6/prodata/1/pubpna/US15_NEW_PUB.seq:*
- 13: /cgn2_6/prodata/1/pubpna/US16_NEW_PUB.seq:*
- 14: /cgn2_6/prodata/1/pubpna/US17_NEW_PUB.seq:*
- 15: /cgn2_6/prodata/1/pubpna/US18_NEW_PUB.seq:*
- 16: /cgn2_6/prodata/1/pubpna/US19_NEW_PUB.seq:*
- 17: /cgn2_6/prodata/1/pubpna/US20_NEW_PUB.seq:*
- 18: /cgn2_6/prodata/1/pubpna/US21_NEW_PUB.seq:*
- 19: /cgn2_6/prodata/1/pubpna/US22_NEW_PUB.seq:*
- 20: /cgn2_6/prodata/1/pubpna/US23_NEW_PUB.seq:*
- 21: /cgn2_6/prodata/1/pubpna/US24_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1763	100.0	2618	11 US-09-867-570-1	Sequence 1, Appl1
2	1727.5	98.0	8622	11 US-09-867-570-3	Sequence 3, Appl1
3	1722.5	97.7	1369	15 US-10-292-798-1273	Sequence 1273, Ap
4	1721	97.6	1400	14 US-10-183-116-30	Sequence 30, Appl
5	1721	97.6	1400	14 US-10-183-116-30	Sequence 673, App
6	1688	95.7	969	9 US-10-225-567A-673	Sequence 19, Appl
7	1688	95.7	969	10 US-09-995-225-19	Sequence 19, Appl
8	1683	95.5	969	10 US-10-401-397A-1	Sequence 1, Appl1
9	1663	94.3	909	15 US-10-237-467-9	Sequence 9, Appl1
10	1557	88.3	909	15 US-10-237-467-9	Sequence 1055, Ap
11	1416	80.3	1369	15 US-10-017-161-1055	Sequence 897, App
12	1416	80.3	1369	14 US-10-292-798-897	Sequence 15, Appl
13	1416	80.3	1370	15 US-10-183-116-15	Sequence 1599, Ap
14	1402.5	79.6	1370	15 US-10-017-161-1599	Sequence 171, App
15	1397	79.2	997	16 US-10-072-012-171	Sequence 1041, Ap
16	1390	78.8	1369	15 US-10-292-798-1041	Sequence 32, Appl
17	1388	78.7	1604	14 US-10-183-116-32	Sequence 688, App
18	1388	78.7	1604	15 US-10-225-567A-688	Sequence 3, Appl1
19	1383	78.4	969	15 US-10-079-384-3	Sequence 19, Appl
20	1383	78.4	969	15 US-10-240-998-3	Sequence 43, Appl
21	1383	78.4	969	15 US-10-321-807-19	Sequence 19, Appl
22	1383	78.4	969	15 US-10-237-467-11	Sequence 17, App
23	1383	78.4	969	16 US-10-343-650A-43	Sequence 173, App
24	1383	78.4	969	17 US-10-321-807-19	Sequence 177, App
25	1383	78.4	969	17 US-10-314-048A-19	Sequence 169, App
26	1383	78.4	969	18 US-10-481-161-1	Sequence 17, Appl
27	1367	77.5	969	9 US-09-995-225-17	Sequence 17, Appl
28	1367	77.5	969	10 US-09-995-225-17	Sequence 17, Appl
29	1367	77.5	969	15 US-10-237-467-3	Sequence 173, App
30	1365	77.4	1005	16 US-10-072-012-173	Sequence 177, App
31	1339.5	75.4	1005	16 US-10-072-012-177	Sequence 169, App
32	1293	73.3	994	16 US-10-072-012-169	Sequence 17, Appl
33	1279	72.5	1997	14 US-10-219-834-7	Sequence 6, Appl1
34	1270	72.0	966	15 US-10-240-998-6	Sequence 9, Appl1
35	1265	71.8	966	15 US-10-240-998-9	Sequence 239, App
36	1227	69.6	793	15 US-10-101-510-239	Sequence 1, Appl1
37	1031.5	58.5	993	17 US-10-467-616-1	Sequence 26, Appl
38	1009.5	57.3	1018	16 US-10-311-673-11	Sequence 11, Appl
39	1009.5	57.3	1163	9 US-09-750-373-11	Sequence 11, Appl
40	1009.5	57.3	1176	18 US-10-481-161-11	Sequence 17, App
41	1009.5	57.3	1300	14 US-10-183-116-17	Sequence 648, App
42	1009.5	57.3	1300	15 US-10-225-567A-648	Sequence 1053, App
43	1009.5	57.3	1393	15 US-10-017-161-1053	Sequence 895, App
44	1009.5	57.3	1393	15 US-10-292-798-895	Sequence 8, Appl1
45	1009.5	57.3	1770	9 US-09-920-068A-8	

ALIGNMENTS

RESULT 1
US-09-867-570-1
Sequence 1, Application US/09867570
Publication No. US20040076951A1
GENERAL INFORMATION:
APPLICANT: WEI, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
FILE REFERENCE: C1000900-CIP
CURRENT APPLICATION NUMBER: US/09/867,570
PRIOR APPLICATION NUMBER: 09/695,045
PRIOR FILING DATE: 2000-10-25
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2618
TYPE: DNA
ORGANISM: Human
US-09-867-570-1

Alignment Scores:
Pred. No.: 1.59e-175 Length: 2618
Score: 1763.00 Matches: 337
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0
US-09-867-570-2 (1-337) x US-09-867-570-1 (1-2618)
QY 1 MetGluSerIySerSerTrpValIleArgLeuGlyPheLeuSerMetAspSerThrIle 20
Db 447 ATGAAACCAATCTCTCATGGGTCTCATCACTGGGTTCTGAGCATGATTCACCAATC 506
QY 21 ProValLeuGlyThrGluLeuThrProIleAenGlyArgGluGluThrProCysTyrLys 40
Db 507 CCAGCTTGGGTACAGAACTGACCAATCAACGAGCGAGGAGAGACTCTTCTCAAG 566
QY 41 GlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsn 60
Db 567 CAGACCTGAGCTTACAGGGGCTGAGTGCATCGTTCCCTGTCCGCTGACAGAAAC 626
QY 61 AlaValIleLeuTrpLeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIle 80
Db 627 GCGGTGTGCTGTGGCTCTGGGCTGCCGATGCGAGAGACGCTGTCTCCATCTACATC 686
QY 81 LeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeu 100
Db 687 CTCAACTGTGTGCGGCGCACTTCTCTTACGAGGCGCACTATATGTTGCGCGTTA 746
QY 101 ArgLeuIleAsnIleArgHisProIleSerIySileLeuSerProValMetThrPhePro 120
Db 747 GCGCTCATCAATATCCGATCCATCCATCTCAAAATCTCAATGTCGTGATGACCTTCC 806
QY 121 TyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeu 140
Db 807 TACTTTATAGGCTTAAGCATGCTGAGCGGCATACAGACCGAGGCTGCTTCATCTC 866
QY 141 TrpProIleTrpTyrHisCysArgArgProArgTyrLeuSerSerValMetCysValLeu 160
Db 867 TGGCCATCTGTGTACCACTGCGCGCCGCCAGATACCTGTCATCGGTGATGTGTC 926
QY 161 LeuTrpAlaLeuSerLeuLeuArgSerIleLeuGluThrMetPheCysAspPheLeuPhe 180
Db 927 CTCTGGGCTGTCTCTGCTGCGAGATCTCGAGTGTGATGTCGTGATCTTCTGTT 966
QY 181 SerGlyAlaAspSerValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuVal 200
Db 987 AGTGTGTGATTTCTGTTGGTGTGAAGTCAATTTTCAATTCATTCATTCGCGTGTG 1046
QY 201 PheLeuCysValValLeuCysGlySerSerLeuValLeuValArgIleLeuCysGly 220
Db 1047 TTTTATATGTGTCTCTGTGTGTGTCAGCTGTGTCGTGTCAGATTTCTCTGTGA 1106
QY 221 SerArgLysMetProLeuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPhe 240
Db 1107 TCCCGAAGATGCCGTGACAGGCTGACGTAACAATCTCTCAACAGCTGTGCTTC 1166
QY 241 LeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAsp 260
Db 1167 CTCCTGTGGGCTGTGCTTGGCATTAGTGGGCTGTGTTTCCAGATCCACCTGTGAT 1226
QY 261 TrpLysValLeuPheCysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSer 280
Db 1227 TGGAAAGCTTATTTGTGATGTGATGATGATTTCCATTTCTGTCCGCTTTAAACAG 1286
QY 281 SerAlaAsnProIleIleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgGln 300
Db 1287 AGTGCACACCCATCATCTTACTTCTGTGGGCTCTCTTAAGGAGGCTCAAAATTAAGCAG 1346
QY 301 AsnLeuLysLeuValLeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGly 320

Db 1347 AACCTGAAGCTGTCTCTCCAGAGGGCTCTGACAGACAGCCCTGAGTGAAGTGA 1406
QY 321 GlyTrpLeuProGlnGluThrLeuGluLeuSerGlySerArgLeuGluGln 337
Db 1407 GGTGTGCTTCTTCAGAGAAACCTGAGGCTGTGGAGAGATGAGTGAACAG 1457
RESULT 2
US-09-867-570-3
; Sequence 3, Application US/09867570
; Publication No. US20040076951A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: CL000900-CIP
; CURRENT APPLICATION NUMBER: US/09/867,570
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 09/695,045
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 8622
; TYPE: DNA
; ORGANISM: Human
US-09-867-570-3
Alignment Scores:
Pred. No.: 5.4e-171 Length: 8622
Score: 1727.50 Matches: 332
Percent Similarity: 96.81% Conservative: 2
Best Local Similarity: 97.23% Mismatches: 2
Query Match: 97.99% Indels: 9
DB: 11 Gaps: 1
US-09-867-570-2 (1-337) x US-09-867-570-3 (1-8622)
QY 2 GluSerLysSerSerTrp-----ValIleArgLeuGly 12
Db 7459 CAGAGATCAAAACAGCTGTGATACATCTGTTCTGTTTCCAGGGTCAATCAGACTGGG 7518
QY 13 PheLeuSerMetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGly 32
Db 7519 TTTCTAGCATGATTCATCAACCATCCAGCTTGGGTACAGAACTACACCAATCAACGA 7578
QY 33 ArgGluGluThrProCysTyrLysGlnThrLeuSerPheThrGlyLeuThrCysIleVal 52
Db 7579 CGTGAAGAGACTCTTGTCTTACAGACAGACCTGAGCTTCAACGGGCTGACGTGATCTGT 7638
QY 53 SerLeuValAlaLeuThrGlyAsnAlaValIleLeuTrpLeuLeuGlyCysArgMetArg 72
Db 7639 TCCCTGTGCGCTGACAGAAACGGGTGTGCTCTGCTCTGCGCTGCGCATGCGC 7698
QY 73 ArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSer 92
Db 7699 AGGAACGCTGTCTTCATCTTCACTCACTGAGCGGCGGAGCTTCTCTTCTTAC 7758
QY 93 GlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerLysIle 112
Db 7759 GGCACATATATATGTTGCGCGTTACGCTCATCAATATCCGATCCCATCCCAAAATC 7818
QY 113 LeuSerProValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSer 132
Db 7819 CTCAGTCTGTGATGACCTTCTTCCCTTATAGGCTTAAGCATCTGAGCGCCATCAGC 7878
QY 133 ThrGluArgCysLeuSerIleLeuTrpProIleTrpTyrHisCysArgArgProArgTyr 152
Db 7879 ACCGAGCGCTGCTGTTCATCTCTGAGGCCATCTGTACCATCTGCCGCCGCCAGATAC 7938
QY 153 LeuSerSerValMetCysValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGlu 172
Db 7939 CTGTATCGGTATGTGTGTCTGTGCTGCGGCTGTCTCTGCTGAGGATCTCTGAG 7998

```

Qy 173 TrpMetPheCybAapPheLeuPheSerGlyAlaAspSerValTrpCysGluThrSerAsp 192
Db 7999 TGAATGTTCTGTGACTTCCTGTTAGTGAGTGCATTCCTGTTGGTGTGAACCGTCAGAT 8058
Qy 193 PheIleThrIleAlaTrpLeuValPheLeuCybValValLeuCysGlySerSerLeuVal 212
Db 8059 TTCATTACAAACGGGAGCGCTGGTTTATTATGATGTTCTCTGTGGGTCCAGCGCTGGTC 8118
Qy 213 LeuLeuValAlaArgIleLeuCybGlySerArgIleMetProLeuThrArgLeuTyrValThr 232
Db 8119 CTGCTGTCAGAGATTCCTCTGTGATCCCGAAGATGCGCTGACCAAGCTGTAGTACC 8178
Qy 233 IleLeuLeuThrValLeuValPheLeuLeuCybGlyLeuProPheGlyIleGlnTrpAla 252
Db 8179 ATCCTCTCACAGCTGCTGCTCTCTCTGTCCTGCTGCGCTGACCTTGGCATTCAGTGGGCC 8238
Qy 253 LeuPheSerArgIleHISLeuAspTrpValLeuPheCybHISValHISLeuValSer 272
Db 8239 CTGTTTCCAGAGATCCAGCTGAGATTGAAAGCTTATTGTCAMTGTGCATCTAGTTTCC 8298
Qy 273 IlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySer 292
Db 8299 ATTTCTGTCCTCGCTCTTAACAGCAGTGCACACCCCATTCATTCTTCTGTGGCTCC 8358
Qy 293 PheArgGlnArgGlnAsnArgGlnAsnLeuValLeuGlnArgAlaLeuGlnAsp 312
Db 8359 TTTAGCAGCGCTCAAAATAGCAGAACCTGAACTGCTTCCAGAGGGCTCTCAGAGAC 8418
Qy 313 ThrProGluValAlaArgGluGlyGlyIleTyrLeuProGlnGluThrLeuGluLeuSerGly 332
Db 8419 ACGCTTAGGATGATGAAGGTGAGGGGTGCTTCTCTCAGAGAAACCTGAGACTGTGGGA 8478
Qy 333 SerArgLeuGluGln 337
Db 8479 AGCAGATTGAGACG 8493

```

RESULT 3

```

US-10-292-798-1273
; Sequence 1273, Application US/10292798
; Publication No. US2003023583A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: AUBURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292, 798
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1273
; LENGTH: 1369
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(1369)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1169)
; US-10-292-798-1273

```

```

Alignment Scores: 1,14e-171 Length: 1369
Score: 1722.50 Matches: 331
Percent Similarity: 96.81% Conservative: 3

```

```

Best Local Similarity: 95.94% Mismatches: 2
Query Match: 97.70% Indels: 9
DB: 15 Gaps: 1
US-09-867-570-2 (1-337) x US-10-292-798-1273 (1-1369) -----ValIleArgLeuGly 12
Qy 2 GluSerIysSerSerTrp-----ValIleArgLeuGly 12
Db 132 CAGAGATCAAAACAGCTGTGATCATCATGCTTCTGTTCCAGAGTCATCAGCTGAGG 191
Qy 13 PheLeuSerMetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGly 32
Db 192 TTTCTGACATGATTTCAACCATCCAGCTTGGGTACAGAACCATCAAGCA 251
Qy 33 ArgGluGluThrProCysTyrIleGlnThrLeuSerPheThrGlyLeuThrCysIleVal 52
Db 252 CGTAGAGAGACTCCTTCTCAAGCAGACCTGAGCTTCAAGGGGCTGAGCGTGCATCGTT 311
Qy 53 SerLeuValAlaLeuThrArgIleAsnAlaValValLeuThrLeuLeuGlyCysArgMetArg 72
Db 312 TCCCTTGTCCGCTGACAGAGAAACGCGTGTGTGCTGTGGCTCTCGGGCTCGCGCATCGC 371
Qy 73 ArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSer 92
Db 372 AGAAGCCTGTCTTCATCTACATCTCAACCTGGTCGCGGCCACATCTCTTCTTACG 431
Qy 93 GlyHISIleIleCysSerProLeuArgLeuIleAsnIleArgHISProIleSerIle 112
Db 432 GGCACATTATATGTTGGCCGTTACGCTCATCAATCCGCCATCCCATCTCCAAATC 491
Qy 113 LeuSerProValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSer 132
Db 492 CTCAGTCTCTGTATGACCTTTCCTTATATAGCCTTAAGCAGATGCTGAGGCGCATCAGC 551
Qy 133 ThrGluArgCysLeuSerIleLeuThrProIleTyrHISCysArgArgProArgTyr 152
Db 552 ACCGAGCGCTGCTGTTCATCTGTGGCCCATCTGGTACCACTGCCGCCCGCCAGATAC 611
Qy 153 LeuSerSerValMetCysValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGlu 172
Db 612 CTGTCAATCAATGATGTGTGCTGTGGGCCCTGTCCCTGCTGCGGAGATCTCGAG 671
Qy 173 TrpMetPheCybAapPheLeuPheSerGlyAlaAspSerValTrpCysGluThrSerAsp 192
Db 672 TGAATGTTCTGTGACTTCCTGTTAGTGAGTGCATTCCTGTTGGTGTGAACCTCGAT 731
Qy 193 PheIleThrIleAlaTrpLeuValPheLeuCybValValLeuCysGlySerSerLeuVal 212
Db 732 TTCATTACAAACGGGAGCGCTGGTTTATTATGATGTTCTCTGTGGGTCCAGCGCTGTC 791
Qy 213 LeuLeuValAlaArgIleLeuCybGlySerArgIleMetProLeuThrArgLeuTyrValThr 232
Db 792 CTGCTGTCAGAGATTCCTCTGTGATCCCGAAGATGCGCTGACCAAGCTGTAGTACC 851
Qy 233 IleLeuLeuThrValLeuValPheLeuLeuCybGlyLeuProPheGlyIleGlnTrpAla 252
Db 852 ATCCTCTCACAGCTGCTGCTCTCTCTGTCCTGCTGCGCTTGGCATTCAGTGGGCC 911
Qy 253 LeuPheSerArgIleHISLeuAspTrpValLeuPheCybHISValHISLeuValSer 272
Db 912 CTGTTTCCAGAGATCCAGCTGAGATTGAAAGCTTATTGTCAMTGTGCATCTAGTTTCC 971
Qy 273 IlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySer 292
Db 972 ATTTCTGTCCTCGCTCTTAACAGCAGTGCACACCCCATTCATTCTTCTGTGGCTCC 1031
Qy 293 PheArgGlnArgGlnAsnArgGlnAsnLeuValLeuGlnArgAlaLeuGlnAsp 312
Db 1032 TTTAGCAGCGCTCAAAATAGCAGAACCTGAACTGCTTCCAGAGGGCTCTCAGAGAC 1091
Qy 313 ThrProGluValAlaArgGluGlyGlyIleTyrLeuProGlnGluThrLeuGluLeuSerGly 332
Db 1092 ACGCTTAGGATGATGAAGGTGAGGGGTGCTTCTCAGAGAAACCTGAGACTGTGGGA 1151

```

```

QY      333 SerArgLeuGluGln 337
DB      1152 AGCAGATTGAGACG 1166

RESULT 4
US-10-183-116-30
; Sequence 30, Application US/10183116
; Publication No. US20030092035A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzhang
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-kyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE, 4C1CPI
; CURRENT APPLICATION NUMBER: US/10/183,116
; CURRENT FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (332)...(1297)
US-10-183-116-30

Alignment Scores:
Pred. No.: 1,7e-171 Length: 1400
Score: 1721.00 Matches: 330
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 97.62% Indels: 0
DB: 14 Gaps: 0

US-09-867-570-2 (1-337) x US-10-183-116-30 (1-1400)
QY      6 SerTPValIleArgLeuGlyPheLeuSerMetArgSerThrIleProValIleGlyThr 25
DB      302 TTCAGGGCTACACAGATGGGGTTCTGAGCATGATTCACCACTCCAGCTGGGATAC 361
QY      26 GluLeuThrProIleAsnGlyArgGluThrProCysEtyrIysGlnThrLeuSerPhe 45
DB      362 GAACGACACCAATCAACGACGACGACGACCTCTTCTGACACAGACCCGAGCTTC 421
QY      46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValIleLeuTyr 65
DB      422 ACGGGGCTGACGTCATCGTTCTTCCTTCGCGCGCTGACAGAAACGGGTTGCTCTGG 481
QY      66 LeuLeuGlyCysArgMetArgArgAsnAlaValSerIleValIleLeuAsnLeuValAla 85
DB      482 CTCCTGGGCTGCCGATGCGACAGAAACGCTGCTCCATCTACACTCCCAACCTGGTCGG 541
QY      86 AlaApphLeuPheLeuSerGlyIleIleCysSerProLeuArgLeuIleAsnIle 105
DB      542 GCCGACTTCTCTTCTTACGGCCACATTAATATGTCGCGCTTACCCCTCATCAATATC 601
QY      106 ArgHisProIleSerIysIleLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
DB      602 CGCCATCCCATCTCCAAATCTCAGTCTGTGATGACCTTTCCTACTTATAGGCTTA 661

```

```

QY      126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTyrProIleTyr 145
DB      662 AGCATGCTAGGCACTACGACACGAGCGCTGCTGCTTCATCTTGGCCCATCTGGTAC 721
QY      146 HisCysArgArgProArgTyrLeuSerSerValMetCysValIleLeuTyrAlaLeuSer 165
DB      722 CACTGCGCGCGCCAGATACCTGTCATCGGTCATGTCGTCTGCTCGGGCCCTGTC 781
QY      166 LeuLeuArgSerIleLeuGluTyrMetPheCysApphLeuPheSerGlyAlaAspSer 185
DB      782 CTGCTGCGAGATCTCTGAGAGGATGTCCTGTCTGCTCTCTGTTAGTGCGCTGATCT 841
QY      186 ValTyrCysGluThrSerApphLeuThrIleAlaTyrLeuValPheLeuCysValVal 205
DB      842 GTTGGTGAGAAACGTCATTCCTTACATTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 901
QY      206 LeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgIleMetPro 225
DB      902 CTCCTGGGCTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961
QY      226 LeuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeu 245
DB      962 CTGACCCAGGCTGTACGTGACCATCTCTCCACAGGCTGCTGCTCTCTCTGTCGCTG 1021
QY      246 ProPheGlyIleGluTyrAlaLeuPheSerArgIleHisLeuAspTyrIleValLeuPhe 265
DB      1022 CCGTTTGCAATTCAGTGGCCCTGTTTTCAGAGATCCACTGATTTGGAATCTTATTT 1081
QY      266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIle 285
DB      1082 TGTGATGTCATCTAGTTTCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1141
QY      286 IleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgGlnAsnLeuVal 305
DB      1142 ATTACTTCTTCTGCTGCTGCTGCTTTCAGGCAAGCTCAAAATAGGCAACCTGAGT 1201
QY      306 LeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGlyTyrLeuProGln 325
DB      1202 CTCGAGAGGCTCTTCAGAGACAGCGCTGAGGTGATGAGAGGAGGAGGCTTCTCAG 1261
QY      326 GluThrLeuGluLeuSerGlySerArgLeuGluGln 337
DB      1262 GAAACCTGAGCTGTCCGGAGACGATTTGAGACG 1297

RESULT 5
US-10-225-567A-673
; Sequence 673, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Lifespan Biociences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 673
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-673

Alignment Scores:
Pred. No.: 1,7e-171 Length: 1400
Score: 1721.00 Matches: 330
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2

```

```

Query Match: 97.62% Indels: 0
DB: 15 Gaps: 0
US-09-867-570-2 (1-337) x US-10-225-567A-673 (1-1400)

Qy 6 SerTrpValIleArgLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThr 25
Db 302 TCACAGGCTACACAGACGCGGGTTTCTGACATGATTCACACATCCAGCTTGGGTACA 361
Qy 26 GluLeuThrProIleAsnGlyArgGluGluThrProCysTyrIlysglnThrLeuSerPhe 45
Db 362 GAACGTACACCAATCAACGAGCGAGAGCTCTTGCTCAACAGACGACCTGAGCTTC 421
Qy 46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValValLeuTrp 65
Db 422 ACCGGGCTGACGGATCGTTTCTTCCCTGTGCGCTGACAGAAACGGGGTGTGCTCTGG 481
Qy 66 LeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAla 85
Db 482 CTCTGGGCTGCGCATGCGGACGAGACGCTGTCTTCATCTACATCTCCACCTGGTGGC 541
Qy 86 AlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleAsnIle 105
Db 542 GCGGACTTCTCTTCTTAGGGGCAATTATATGTTGCGCGTTAGCGCTCATCAATATC 601
Qy 106 ArgHisProIleSerIlyIleLeuSerProValMetThrPheProIlyPheIleGlyLeu 125
Db 602 CGGCATCCCATCTCCAAATCTCATGCTGTGATGATGATCTTCTTCTTATAGGCTTA 661
Qy 126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTrpProIleTrpTyr 145
Db 662 AGCATGCTGAGCGCATACAGACGAGCGCTGTGCTTCATCTGAGCCATCTGGTAC 721
Qy 146 HisCysArgArgProArgTyrLeuSerSerValMetCysValLeuLeuTrpAlaLeuSer 165
Db 722 CACTGCGCGCGCCCAATACCTGTCATCGGTATGATGTCGTCGCTGGGCTCTGTC 781
Qy 166 LeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeuPheSerGlyAlaAspSer 185
Db 782 CTCTCGCGGAGTATCCGAGAGTGTCTGTGATCTCTGTTTACTGTGCTGATCTCT 841
Qy 186 ValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuValPheLeuCysValVal 205
Db 842 GTTGGTGTAAACGTACATTCATTACATGCGGGCTGTTTATATGTGTGTT 901
Qy 206 LeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgLysMetPro 225
Db 902 CTCTGGGGTCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961
Qy 226 LeuThrArgLeuTyrValIleThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeu 245
Db 962 CTACACAGGCTGTACGACCATCTCTCAGAGTGTGCTCTCTCTGCGGCTG 1021
Qy 246 ProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAspTrpIlyValLeuPhe 265
Db 1022 CCTTTGGCATTCAGTGGGCTCTGTTTTCAGAGATCCACCTGATGGAAGCTTAATTT 1081
Qy 266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIle 285
Db 1082 TGTCAATGTGATCTAGTTTCATTTCTCTGCTCCCTTTTAAACAGCATGCGCAACCCATC 1141
Qy 286 IleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgGlnAsnLeuValLeuVal 305
Db 1142 ATTACTTCTTCTGCGGCTCTCTTAGGCAAGCTCAAAATAGGCGAAACCTGAAAGCTGTT 1201
Qy 306 LeuGluArgAlaLeuGlnAspTrpProGluValAspGluGlyGlyIlyTrpLeuProGln 325
Db 1202 CTCACAGGAGCTGTGACGACGACCTGAGGTGATGAGGTGAGGTGAGGTGAGGTGAG 1261
Qy 326 GluThrLeuGluLeuSerGlySerArgLeuGluGln 337
Db 1262 GAAACCTGAGCTGTGCGGAGACAGATTGAGAG 1297

```

* RESULT 6

```

US-09-995-225-19
; Sequence 19, Application US/09995225
; Publication No. US20020193584A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huong T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Pryde, Cameron
; TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human G
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: No. US20020193584A1el Sequence
US-09-995-225-19

```

Alignment Scores:

```

Pred. No.: 2,99e-168 Length: 969
Score: 1688.00 Matches: 322
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.75% Indels: 0
DB: 9 Gaps: 0

```

US-09-867-570-2 (1-337) x US-09-995-225-19 (1-969)

```

Qy 16 MetAspSerThrIleProValIleGlyThrGluLeuThrProIleAsnGlyArgGluGln 35
Db 1 ATGAGATTCAACCAATCCAGCTTGGGTACAGAACTGACACCAATCAACGAGCTGAGAG 60
Qy 36 ThrProCysTyrIlysglnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
Db 61 ACTCTTGTCTACACAGACAGACCTGAGCTTCAAGGGCTACGTCATCTTCTGTC 120
Qy 56 AlaLeuThrGlyAsnAlaValValLeuTrpLeuGlyCysArgMetArgAsnAla 75
Db 121 GCGCTACAGGAACCGGGTGTGCTCTGCTCTGCGGCTGCGGCAATGCGAGAACGCT 180
Qy 76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuSerGlyHisIle 95

```

```

Db      181  GTCTCATCTACATCTGACCTGCTCGGCGGCGAATTCTCTCTTAGCGCCACATT 240
Qy      96  TLeCySerProLeuAryLeuIleasnIleAghIspProIleSerIyIleLeuSerPro 115
Db      241  ATATGTTGCGCGTTAGCGCTCATCAATATCCCGCATCCCATCTCCAAAATCTCAGTCT 300
Qy      116  ValMetThrPheProTyPheIleGlyLeuSerMetLeuSerAilIleSerThrgIuArg 135
Db      301  GTGATGACCTTCTCTTATTAAGCCTTAAGCACTGACAGCCATCCATCCAGACCGCC 360
Qy      136  CyLeuSerIleLeuTrpProIleTrpTyRHisCyAArgArgProArgTyRLeuSerSer 155
Db      361  TGCGGTGCATCTCTGCGCCCATCTGTFACCACTGCGCGCGCCGAGATACCTGCATCG 420
Qy      156  ValMetCyValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGluTrpMetPhe 175
Db      421  GTCATGTGTGCTCTGCTCTGCGCCCTGTCCTCCGCGGAGATATCTGAGAGGAGTTC 480
Qy      176  CyAspPheLeuPheSerGlyAlaAspSerValTrpCyGluThrSerAspPheIleThr 195
Db      481  TGTGATCTCTGTTAGTGTGAGTGTGATTTGTTGTGTGTAACGTCAGATTTCAATTACA 540
Qy      196  IleAlaTrpLeuValPheLeuCyValValLeuCyGlySerSerLeuValLeuVal 215
Db      541  ATCGGTGCGTGGTGTCTTTTATGTGTGTCTCTGTGTGCTGACAGCTGCTGCTGCTGTC 600
Qy      216  ArgIleLeuCyGlySerArgIyMetProLeuThrArgLeuTyTAlThrIleLeuLeu 235
Db      601  AGGATTCCTCTGTGTGATCCCGGAAGATGCGGCTGACAGAGCTGATACCATCTCTCTC 660
Qy      236  ThrValLeuValPheLeuLeuCyGlyLeuProPheGlyIleGluTrpAlaLeuPheSer 255
Db      661  ACAGTGTGCTGTCTCTCTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTTC 720
Qy      256  ArgIleHisLeuAspTrpIyValLeuPheCyHisValHisLeuValSerIlePheLeu 275
Db      721  AGGATCCACCTGTGATGGAAAGCTTATTTGTGTCATGTGATCAATTTCATTTTCCGTC 780
Qy      276  SerAlaLeuAsnSerSerIleAsnProIleIleTyRPhPheValIlySerPheArgGlu 295
Db      781  TCCGCTCTTAAAGAGAGTGCACCAACCCATCATTTACTTCTCTGCGGCTCTTTAGGCG 840
Qy      296  ArgGluAsnArgGluAsnLeuIyLeuValIleGluAlaArgAlaLeuGluInAspTrpProGlu 315
Db      841  CGTCAAAATAGCAGAAACCTGAAAGCTGTCTTCCAGAGGCTTGCAGAGCACGCTGAG 900
Qy      316  ValAspGluGlyGlyTyRLeuProGluGluThrLeuGluLeuSerGlySerArgLeu 335
Db      901  GTGATGAAAGTGTGAGGCTGCTTCTCAGAAACCTGGAAGCTGTGCGGAGCAGATTG 960
Qy      336  GluGlu 337
Db      961  GAGCAG 966

```

RESULT 7

```

US-09-995-225-19
; Sequence 19, Application US/0995225
; Publication No. US20030139588A9
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huang T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Fried, Cameron
; TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human C
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13

```

```

; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: NO. US20030139588A9el Sequence
US-09-995-225-19

Alignment Scores:
Pred. No.: 2,99e-168 Length: 969
Score: 1688.00 Matches: 322
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.75% Indels: 0
DB: 10 Gaps: 0

US-09-867-570-2 (1-337) x US-09-995-225-19 (1-969)
Qy      16  MetAspSerThrIleProValLeuGlyThrGluLeuTrpProIleAsnGlyArgGluGlu 35
Db      1  ATGGATTCAACATCCCACTGTGGGTACAGAACCTGACCAATCAAGACGTAGAG 60
Qy      36  ThrProCyTyRtyGluInThrLeuSerPheThrGlyLeuThrCyAlIleValSerLeuVal 55
Db      61  ACTCTTGCTTACAGACAGACCTGAGCTTACAGGGGCTGACGTGATGCTTCCCTTGTTC 120
Qy      56  AlaLeuThrGlyAsnAlaValValLeuTrpLeuLeuGlyCyAArgMetArgArgAsnAla 75
Db      121  GGGCTGACAGAAAGCGGTTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
Qy      76  ValSerIleTyRtleuAsnLeuValAlaIleAspPheLeuPheLeuSerGlyHisIle 95
Db      181  GTCTCATCTACATCTCTCAACCTGTGCGGCGGCGGCACTTCTCTTACGCGGCACATT 240
Qy      96  TLeCySerProLeuAryLeuIleasnIleAghIspProIleSerIyIleLeuSerPro 115
Db      241  ATATGTTGCGCGTTAGCGCTCATCAATATCCCGCATCCCATCTCCAAAATCTCAGTCT 300
Qy      116  ValMetThrPheProTyPheIleGlyLeuSerMetLeuSerAilIleSerThrgIuArg 135
Db      301  GTGATGACCTTCTCTTATTAAGCCTTAAGCACTGACAGCCATCCATCCAGACCGCC 360
Qy      136  CyLeuSerIleLeuTrpProIleTrpTyRHisCyAArgArgProArgTyRLeuSerSer 155
Db      361  TGCGGTGCATCTCTGCGCCCATCTGTFACCACTGCGCGCGCCGAGATACCTGCATCG 420
Qy      156  ValMetCyValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGluTrpMetPhe 175
Db      421  GTCATGTGTGCTCTGCTCTGCGCCCTGTCCTCCGCGGAGATATCTGAGAGGAGTTC 480
Qy      176  CyAspPheLeuPheSerGlyAlaAspSerValTrpCyGluThrSerAspPheIleThr 195

```


Db	481	TGTGACTTCCTGTTAGTGGCTGTAACTCTGTTGGTGTGAACGTCAGATTTCATTACA	540
Qy	136	IIeaIaTPrLeuValPheIeuCYsValIalleuCYsGlySerSerLeuValIleuLeuVal	215
Db	541	ATCGGCTGAGCTGGTTTTTTTATATGTGTGGTCTCTGTGGGTCACGCTGGCTGCTGGTC	600
Qy	216	ArgIleIeuCYsGlySerArgIysMetProIeuThrArgIeuIyrAlaThrIleIeuIeu	235
Db	601	AGGATTTCTGTGTAAATCCGGAAGATGCGCGCTGACCAAGGCTGTACGTACATCTCTC	660
Qy	236	ThrValIleuValPheIeuIeuCYsGlyIleuProPheGlyIleGIInTrpAlaIeuPheSer	255
Db	661	ACAGTGGCTGGTCTTCTCTCTCTGTGGCCGTGCCCTTTGGCATTCAGTGGGCGCTGTTCC	720
Qy	256	ArgIIeHIIleuIaSPTrIyValIeuPheCYsHIIValIleuValSerIIePheIeu	275
Db	721	AGGATCCACCTCGGATTGGAAAGCTTAATTTTGTGCATGTGCATCTAATTTCATTCTCG	780
Qy	276	SerAlaIeuAAsrSerIalaAsnProIleIeIyPhePheValGlySerPheAArgIn	295
Db	781	TCCGCTCTTAACAGCAGTGCACAAACCCATCATTTACTTCTTCGTGGGCTCTTTAGGCAG	840
Qy	296	ArgIAsnAsnArgGlnAsnIleuIysLeuValIleuGlnArgAlaIeuGlnAsnTrpGlu	315
Db	841	CCTCAAAATAGGCACAGAACTGAAAGCTGGTCTTCACAGAGGCTTGCAGAGACAGCCTGAG	900
Qy	316	ValAspGluGlyGlyGlyIyTrIleuProGlnIyThrIleuGluIeuSerGlySerArgIeu	335
Db	901	GTCGTATGAAGGTGGAGGCTGCTTCTCTCAGGAACCTCGAGCTGTCTGGAGACAGATTG	960
Qy	336	GIuGln 337	
Db	961	GAGCAG 966	

```

RESULT 8
US-10-401-397A-1
; Sequence 1, Application US/10401397A
; Publication No. US20030212001A1
; GENERAL INFORMATION:
; APPLICANT: Peri, Krishna G.
; APPLICANT: Moffett, Serge
; APPLICANT: Abran, Daniel
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR PREVENTION AND TREATMENT OF ELEVATED
; TITLE OF INVENTION: INTRAOCULAR PRESSURE AND RELATED CONDITIONS
; FILE REFERENCE: 4518/1M674US1
; CURRENT APPLICATION NUMBER: US/10/401.397A
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/367,513
; PRIOR FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(969)
; OTHER INFORMATION: coding sequence for polypeptide
; US-10-401-397A-1

```

Alignment Scores:	
Pred. No.:	1.0e-167
Score:	1683.00
Percent Similarity:	100.00%
Best Local Similarity:	99.69%
Query Match:	95.46%
DB:	15
Length:	965
Matches:	322
Conservative:	1
Mismatches:	0
Indels:	0
Gaps:	0

US-09-867-570-2 (1-337) x US-10-401-397A-1 (1-969)

16 MetaspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGlyArgGluGlu 35

Db	1	ATGATGTTCAACCATCCAGTCTTGGGTTCAGAACTGACCAATCAACGGAAGCTGAGAG	60
QY	36	ThrProCysTrpLysGlnThrIleuSerPheMetArgIleuThrCysIleValSerLeuVal	55
Db	61	ACTCCTTGCTGCACACAGACCTCTGAGCTTCAAGGGAGTCACTGCATCTCTTCCCTTGTC	120
QY	56	AlaIleuThrGlyAsnAlaValValIleuTrpLeuLeuGlyCysArgMetArgArgAsnAla	75
Db	121	GGCTGACAGAGAGACGGGTTGTGTCTGGCTCTTGCGCTCGCCGACATGGCAGAACGCT	180
QY	76	ValSerIleTrpIleLeuAsnLeuValAlaAlaSerPheLeuPheLeuSerGlyHisIle	95
Db	181	GGTCCATCTACATCTCTCAACTGGATGGGGCCGACCTTCTCTTCCCTTAGCGGCACATT	240
QY	96	IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerTrpIleLeuSerPro	115
Db	241	ATATGTGGCCGTACCGCTCATCAATATACCCCATCCATCTCCAAATATCCATCTCT	300
QY	116	ValMetCysPheProTrpPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg	135
Db	301	GTGATGACCTTTCCCTTACTTTATAGGCTTAACGATGATGAGGCCATCAGCACCGAGCGC	360
QY	136	CysLeuSerIleLeuTrpProIleTrpTrpHisCysArgArgProArgTrpLeuSerSer	155
Db	361	TGCCGTGCTCATCTCTGGGCCATCTGGTACCATCGCCGCCGACGATACCTGTATCG	420
QY	156	ValMetCysValLeuLeuTrpAlaLeuSerIleLeuLeuArgSerIleLeuGluTrpMetPhe	175
Db	421	GTCAATGTGTCTCTCTCTGGGCCCTGTGCTCTGCTGGAGATGATCTTGAGATGCTTC	480
QY	176	CysAspPheLeuPheSerGlyAlaAspSerValTrpCysGluTrpSerAspPheIleTrp	195
Db	481	TGTGACTTCTCTGTAGTGTGCTGATTTCTGTTGGTGGAAACGTCAGATTTCAATTACA	540
QY	196	IleAlaTrpLeuValPheLeuCysValValLeuCysGlySerSerIleuValLeuVal	215
Db	541	ATCGGGTGGCTGGTTTTTATGTGTGTCTCTGTGGGTTCAGCGCTGTCTCGTGGTTC	600
QY	216	ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTrpValThrIleLeuLeu	235
Db	601	AGGATTTCTGTGGATCCCGAAGATGCCGTGACCAAGGCTGATCACTGACATCTCTTC	660
QY	236	ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSer	255
Db	661	ACAGGTGAGTCTTCTCTCTCTGTGGGCCCGCTTTGGCATTCAGGGGCGCTGTTTCC	720
QY	256	ArgIleHisIleAspTrpLysValLeuPheCysHisValHisIleuValSerIlePheLeu	275
Db	721	AGGATTCACCTGGATTTGAAAGTCTTATTTGTCAAGTGCATCTGATTTTCATTTTCCG	780
QY	276	SerAlaLeuAsnSerSerAlaAsnProIleIleTrpPhePheValGlySerPheArgGln	295
Db	781	TCCGCTCTTAAACAGACAGTGGCCACCCCATCTTACTTCTGTGGGCTCTTTAGGCAG	840
QY	296	ArgGlnAsnAlaArgGlnAsnLeuLysLeuValLeuGlnArgAlaLeuGlnAspThrProGlu	315
Db	841	CGTCAAAATAGGCACAAACCTGAGGTGTCTTCCAGAGGGCTCTGACGACACGCTTGAG	900
QY	316	ValAspGluGlyGlyGlyTrpLeuProGlnGluTrpLeuGluLeuSerGlySerArgLeu	335
Db	901	GTGATGTAAGGTGAGGGGTGCTTCTCAAGAAACCTTGAGACTGTGGGAACAGATTG	960

RESULT 9
US-10-391-074-1
; Sequence 1, Application US/10391074
; Publication No. US20040038345A1
; GENERAL INFORMATION:
; APPLICANT: Nehls, Michael

APPLICANT: Mattler, Frank
 TITLE OF INVENTION: No. US20040038345A1el Human Seven-Transmembrane Receptors
 FILE REFERENCE: 7705.0008-00-000
 CURRENT APPLICATION NUMBER: US/10/391,074
 CURRENT FILING DATE: 2003-03-17
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 1
 LENGTH: 969
 TYPE: DNA
 ORGANISM: Homo sapien
 US-10-391-074-1

Alignment Scores:

Pred. No.:	1,31e-165	Length:	969
Score:	1663.00	Matches:	320
Percent Similarity:	99.38%	Conservative:	0
Best Local Similarity:	99.38%	Mismatches:	2
Query Match:	94.33%	Indels:	0
DB:	16	Gaps:	0

US-09-867-570-2 (1-337) x US-10-391-074-1 (1-969)

```

QY 16 MetaspSerThrIleProValIleuGlyThrgIleuThrProIleAsnGlyArgGluGln 35
DB 1 ATGGATTCAACATCCAGTCTTGGGTACAGAACTGACCAATCAACGAGAGTGAGAG 60
QY 36 ThrProCyTyIleGlnThrlleuSerPheThrgIleuThrcysIleValSerIleuVal 55
DB 61 ACTCTTCTTCAACACAGACCTGAGCTTCAACGAGGCTGACGAGCTGCTTCTTCTGTC 120
QY 56 AlaIleuThrgIlyAsnAlaValIleuThrPleuIleuGlyCysArgIleuArgAsnAla 75
DB 121 GGGCTGACAGGAAAGCGGTGTGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 76 ValSerIleTyIleuAsnIleuValAlaAlaAspPheIleuSerGlyIleIle 95
DB 181 GTCTCACTACATCTCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 96 IleCySerProIleuThrgIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 115
DB 241 ATACCTTGGCCCTTACGCTCTCAATCAATCCGCTCAATCCGCTCAATCCGCTCAATCC 300
QY 116 ValMetThrPheProTyIlePheIleGlyIleuSerMetIleuSerAlaIleSerThrgIuArg 135
DB 301 GTGATGACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 360
QY 136 CysIleuSerIleuIleuThrProIleTyIleIleuIleuIleuIleuIleuIleuIleu 155
DB 361 TGGCTGCTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 156 ValMetCysValIleuIleuThrPalaIleuSerIleuIleuIleuIleuIleuIleu 175
DB 421 GTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 176 CysAspPheIleuPheSerGlyAlaAspSerValIleuIleuIleuIleuIleuIleu 195
DB 481 TGTGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540
QY 196 IleAlaIleuValPheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 215
DB 541 ATCGGTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
QY 216 ArgIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 235
DB 601 AGGATTTCTTCTGAGATCCCGGAGAGATCCCGCTGACCAAGGCTTACCGGATCCCTCC 660
QY 236 ThrValIleuValPheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 255
DB 661 ACAAGTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 720
QY 256 ArgIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 275

```

```

DB 721 AGGATCCACCTGATGGAAAGTCTTATTTTGTATGATGATCACTAGTTCCATTTCTCTG 780
QY 276 SerAlaIleuAspSerPheIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 295
DB 781 TCCGCTCTTAAACAGAGTGCACCCCACTTATCTTCTTCTGAGGCTCTTCTTCTTCTTCT 840
QY 296 ArgGlnAsnArgGlnAsnIleuIleuValIleuIleuIleuIleuIleuIleuIleu 315
DB 841 CGTCAAAATAGGCAACACTGAGCTGCTTCTCCAGAGGCTCTGACAGACACGCTGAG 900
QY 316 ValAspGluGlyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 335
DB 901 GTGATGAAAGTGGAGGAGGAGCTTCTCAGGAAACCTGAGCTGTCTGGAAGAGATTG 960
QY 336 GluGln 337
DB 961 GAGCAG 966

```

RESULT 10

US-10-237-467-9
 Sequence 9, Application US/10237467
 Publication No. US20030186324A1

GENERAL INFORMATION:

APPLICANT: Liao, JiaYu
 APPLICANT: Gray, Nathanael S.
 APPLICANT: Caldwell, Jeremy C.
 APPLICANT: Schultz, Peter G.

APPLICANT: IRM LLC
 TITLE OF INVENTION: Sensory Neuron Receptors

FILE REFERENCE: 021288-001300US

CURRENT APPLICATION NUMBER: US/10/237,467

CURRENT FILING DATE: 2003-01-14

PRIOR APPLICATION NUMBER: US 60/317,879

PRIOR FILING DATE: 2001-09-07

NUMBER OF SEQ ID NOS: 18

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 9

LENGTH: 909

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: dorsal root ganglia G-protein coupled receptor (GPCR)

US-10-237-467-9

Alignment Scores:

Pred. No.:	1,88e-154	Length:	909
Score:	1557.00 <td>Matches:</td> <td>301 </td>	Matches:	301
Percent Similarity:	93.79% <td>Conservative:</td> <td>1 </td>	Conservative:	1
Best Local Similarity:	93.48% <td>Mismatches:</td> <td>20 </td>	Mismatches:	20
Query Match:	88.32% <td>Indels:</td> <td>1 </td>	Indels:	1
DB:	15	Gaps:	1

US-09-867-570-2 (1-337) x US-10-237-467-9 (1-909)

```

QY 16 MetaspSerThrIleProValIleuGlyThrgIleuThrProIleAsnGlyArgGluGln 35
DB 1 ATGGATTCAACATCCAGTCTTGGGTACAGAACTGACCAATCAACGAGAGTGAGAG 60
QY 36 ThrProCyTyIleGlnThrlleuSerPheThrgIleuThrcysIleValSerIleuVal 55
DB 60 -----
QY 56 AlaIleuThrgIlyAsnAlaValIleuThrPleuIleuGlyCysArgIleuArgAsnAla 75
DB 61 GGGCTGACAGGAAAGCGGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 76 ValSerIleTyIleuAsnIleuValAlaAlaAspPheIleuPheIleuSerGlyIleIle 95
DB 121 GTTCACTTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 181 IleCySerProIleuThrgIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 200

```

```

Db      181 ATATGTTGGCGCTTACGCGCTCATCATATCCGCCATCCCATCTCCAAATCTCATGTCCT 240
Qy      116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
Db      241 GTGATGACCTTCCCTCACTTATAGGCGCTAAGCATGTCAGACGCCCATCAGACCGAGCGC 300
Qy      136 CysLeuSerIleLeuTrpProIleTrpTyrHisGlySerArgProArgTyrLeuSerSer 155
Db      301 TGGCTGTCAATCCGTGGCCCATCTGTACACACGCGGCCGCCAGATACCTTCATCA 360
Qy      156 ValMetCysValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGluTrpMetPhe 175
Db      361 GTATGTGTGTCCGCTCTGGGCCCTGTCCCTGCTGGCGAATACCTCGAGTGCATGTC 420
Qy      176 CysAspPheLeuPheSerGlyAlaAspSerValTrpCysGluTrpSerAspPheIleThr 195
Db      421 TGTACTCTCCCTGTTAGTGTGTAATCTGTTGTGTGAACGTCAAGATTTCATTACA 480
Qy      196 IleAlaTrpLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuVal 215
Db      481 ATGCGGTGGCTGTTTATTTATGTTGTTCTGTGTGGTCCAGCTGGTCTGTGCTGCTG 540
Qy      216 ArgIleLeuCysGlySerArgGlyMetProLeuThrArgLeuTyrValThrIleLeuLeu 235
Db      541 AGGATTCTGTGTGATCCCGAAGATGCCGTGACAGGCTGTACGTGACCATCTCTC 600
Qy      236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSer 255
Db      601 ACAGTGTGGTCTTCTCTCTCTGTGGCTGCTGCTGGCATTCAGTGGGCCCTGTTTCC 660
Qy      256 ArgIleIleLeuAspTrpPheValLeuPheCysHisValHisLeuValSerIlePheLeu 275
Db      661 AGGATTCACCTGATTGTGAAGTCTTATTTGTATGTCATGTGATTCATGATTTTCTG 720
Qy      276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGln 295
Db      721 TCCGCTCTTAACAGCATGTGCCAACCCATCTTACTTCTTCGTGGGCTCTTTAGGCG 780
Qy      296 ArgGlnAsnArgGlnAsnLeuLeuValLeuGlnArgAlaLeuGlnAspThrProGlu 315
Db      781 CGTCAAAATAGCGAAGACGTGAAGCTGATCTCTCAGAGGGCTGTGAGGACACCGCTGAG 840
Qy      316 ValAspGlnGlyGlyTyrTrpLeuProGlnGluThrLeuGluLeuSerGlySerArgLeu 335
Db      841 GTGATGTAAGGTGAGGGGTGTCTTCACAGAAACCTTGAGCTGTGGGAAGCAGATTG 900
Qy      336 GluGln 337
Db      901 GAGCAG 906

```

```

RESULT 11
US-10-017-161-1055
; Sequence 1055, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIHO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABEYARATNE, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017, 161
; PRIOR FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1055
; LENGTH: 1369
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: source

```

```

; LOCATION: (1)..(1369)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1169)
US-10-017-161-1055

Alignment Scores:
Pred. No.: 2,73e-139 Length: 1369
Score: 1416.00 Matches: 277
Percent Similarity: 88.25% Conservative: 16
Best Local Similarity: 83.43% Mismatches: 39
Query Match: 80.32% Indels: 0
DB: 15 Gaps: 0

US-09-867-570-2 (1-337) x US-10-017-161-1055 (1-1369)
Qy      6 SerTrpValIleArgLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThr 25
Db      171 TCCAGGGTACACAGACTAAGGTTTCTAGCATGATCCACCATCTCACTTGACACA 230
Qy      26 GluLeuThrProIleAsnGlyArgGluGluThrProCysTyrGlyGlnThrLeuSerPhe 45
Db      231 GAACGTACACCAATCAACGGAAGTGAAGAGCTCTTGTCTACAGACAGACCTTGAGCTTC 290
Qy      46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValValLeuTrp 65
Db      291 ACGGTGTGACGTGCATCTGTTCCCTGTGCGGCTGACAGGAAACGAGTGTCTCTG 350
Qy      66 LeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAla 85
Db      351 CTCTGGGCTGCGCCATGCGGAGAGAGCTTCTTCATCTACATCTCACTTCACTTGCGCGCA 410
Qy      86 AlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleAsnIle 105
Db      411 GCAAGCTTCCTCTCTCAGAGGCGGCTTATATATTCCTGTTAAGCTTCATCAGATC 470
Qy      106 ArgHisProIleSerIleValLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
Db      471 CCCCATACCAATCTTAAATCTCTATCTGTGTATGTTTCTCTCTTGCAGGCGCTG 530
Qy      126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTrpProIleTrpTyr 145
Db      531 AGCTTTGTAGTGGCGGAGACCGAGCGCTGCTGTCGCTGTGGCCCATGTGTAC 590
Qy      531 ArgHisProIleSerIleValLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
Db      146 HisCysArgArgProArgTyrLeuSerSerValMetCysValLeuLeuTrpAlaLeuSer 165
Qy      591 CGGTGCCACCGCCCAACACACTGTCAAGCGGTGTGTGTCCTGCTGGGCCCTGTGCC 650
Db      166 LeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeuPheSerGlyAlaAspSer 185
Qy      651 CTGCTGCGAGACATCTCGAGATGTATGTGGCTTCTCTGTTCACTGTGGCTGATTC 710
Db      186 ValTrpCysGluTrpSerAspPheIleThrIleAlaTrpLeuValPheLeuCysValVal 205
Qy      711 GCTTGTGTCAACATCAATGATTCATCAACAGCTGCGGTGATTTTATGTGTGTT 770
Db      206 LeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgGlyMetPro 225
Qy      771 CTCTGTGGGTCCAGCGCTGCTGCTGCTGATGATTCCTGTGATCCCGGAAGATACCG 830
Db      226 LeuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeu 245
Qy      831 CTGACCAAGCGCTGACGACCATCTCTCAAGTACTGTGCTCTCTCTGTGGGCTG 890
Db      246 ProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAspTrpValLeuPhe 265
Qy      891 CCTTGGCATTCAGATTCTTCATTTTATTTATGATCCAGTGAAGAGGATTTATTT 950
Db      266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIle 285
Qy      951 TGTCAATTCATCAGTTCTTATTTCTGTGCGCTTCAACAGCAAGTGCACCACTCCATC 1010
Db      286 IleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgGlnAsnLeuVal 305

```

Db 1011 ATTACTCTTCGTGGGCTCTTTCAGGAGCGCTCAAAATAGCAGAACTGAAAGCTGCTT 1070
Qy 306 LeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGlyTyrTrpLeuProGln 325
Db 1071 CTCAGAGGGCTCTTCAGGAGCGCTGTGAGTGATGAAGTGAGGAGGCTTCTCTGAG 1130
Qy 326 GluThrLeuGluLeuSerGlySerArgLeuGluGln 337
Db 1131 GAAATCTGTGAGCTGTGGGAAAGCAGATTGAGCAG 1166

RESULT 12

US-10-292-798-897
; Sequence 897, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: AIBURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRI-PHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 897
; LENGTH: 1369
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(1369)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(1169)
; US-10-292-798-897

Alignment Scores:

Pred. No.: 2,736-139 Length: 1369
Score: 1416.00 Matches: 277
Percent Similarity: 88.25% Conservative: 16
Best Local Similarity: 83.43% Mismatches: 39
Query Match: 80.32% Indels: 0
DB: 15 Gaps: 0

US-09-867-570-2 (1-337) x US-10-292-798-897 (1-1369)

Qy 6 SerTrpValIleArgLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThr 25
Db 171 TCCAGGGTCACGAGCTAGGTTTCTGAGCATGATCAACATCTCAACCTTGACACACA 230
Qy 26 GluLeuThrProIleAsnGlyArgGluGluThrProGlyTyrIleGlnThrLeuSerPhe 45
Db 231 GAACGACACCAATCAACGGAACCTGAGAGACTCTTGTCAAGAGACCTTGACCTC 290
Qy 46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValAlaLeuTrp 65
Db 291 ACGGCTGACGCTGATCTTTCCTTTCGAGCTTACAGAGAAACGAGTTGCTCTG 350
Qy 66 LeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAla 85
Db 351 CTCCTGGGCTCCGATGCGAGGAGAACGCTTCTCATCTACATCTCAACTTGCGCGCA 410
Qy 86 AlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleAsnIle 105
Db 411 GCAAGCTTCCTCTCTCTGAGCGCGCCTTATATATATCTCTGTTAAGCTTCATCAGTATC 470

Qy 106 ArgHisProIleSerIleIleLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
Db 471 CCCCATACCATCTCTAAATCTCTATCTGTGATGATGATTTTCTTACTTTCGAGGCTG 530
Qy 126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTrpProIleTrpTyr 145
Db 531 AGCTTCTGAGTCCGCTGAGCAGCAGCGCTGCTCCGCTCTGCGCCATCTGAGTAC 590
Qy 146 HisCysArgArgProArgTyrIleLeuSerSerValMetCysValIleLeuThrAlaLeuSer 165
Db 591 CGCTGCCACCGCCCAACACCTGTACAGGAGTGTGTGTCTCTGCGCCCTGCTCC 650
Qy 166 LeuLeuArgSerIleLeuGluThrPheMetPheCysAspPheLeuPheSerGlyAlaAspSer 185
Db 651 CTGCTGCGGAGCATCTCTGAGATGATTTATGTGCTTCTCTTCAAGTGTGCTGATCT 710
Qy 186 ValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuValPheLeuCysValVal 205
Db 711 GCTTGGTGTCAAAACATCATGATTTCAACAGCGGTGCTGATTTTATGTGTGTT 770
Qy 206 LeuCysGlySerSerIleValLeuLeuValArgIleLeuCysGlySerArgIleMetPro 225
Db 771 CTCTGGGTTCAGCGCTGTCTGTGATCAGAAATCTCTGTGATCCCGGAAGATACCG 830
Qy 226 LeuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeu 245
Db 831 CTGACCAAGCTGTACGTGACATCTCTGCTCAAGATCTGTCTCTCTCTGTGCTG 890
Qy 246 PropGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAspTrpIleValLeuPhe 265
Db 891 CCTTGGCATTCAGTTTCTCTATTTTATGATCCAGTGCAGCAGGAGAGTCTTATTT 950
Qy 266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAspSerAlaAspProIle 285
Db 951 TGTCAATGATCATATATTTCTATTTCTGTCCGCTTTAAGAGAGATGCCAACCCTTC 1010
Qy 286 IleTyrPhePheValGlySerPheArgGlnArgGlnAspArgGlnAsnLeuLeuVal 305
Db 1011 ATTACTTCTTGTGGGCTCTTTCAGGAGCGCTCAAAATAGCAGAACTGAAAGCTGCTT 1070
Qy 306 LeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGlyTyrTrpLeuProGln 325
Db 1071 CTCAGAGGGCTCTTCAGGAGCGCTGTGAGTGATGAAGTGAGGAGGCTTCTCTGAG 1130
Qy 326 GluThrLeuGluLeuSerGlySerArgLeuGluGln 337
Db 1131 GAAATCTGTGAGCTGTGGGAAAGCAGATTGAGCAG 1166
RESULT 13
US-10-183-116-15
; Sequence 15, Application US/10183116
; Publication No. US20030092035A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzong
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-kyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE, 4C1CPI
; CURRENT APPLICATION NUMBER: US/10/183,116
; CURRENT FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 109

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 15
 LENGTH: 2040
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURES:
 NAME/KEY: CDS
 LOCATION: (328)...(1293)
 US-10-183-116-15

Alignment Scores:

Pred. No.:	4.98e-139	Length:	2040
Score:	1416.00	Matches:	277
Percent Similarity:	88.25%	Conservative:	16
Best Local Similarity:	83.43%	Mismatches:	39
Query Match:	80.32%	Indels:	0
	14	Gaps:	0

US-09-867-570-2 (1-337) x US-10-183-116-15 (1-2040)

```

Qy 6 SerTrpValIleArgLeuGlyPheLeuSerMetAspSerThrIleProValIleuGlyThr 25
Db 298 TCCAGGAGTACCAAGATAGAGGTTCTGAGCATGATCCAAACCATCTGACACACA 357
Qy 26 GluLeuThrProIleAsnGlyArgGluGluThrProCysTyrIleGlnThrLeuSerPhe 45
Db 358 GAATGACACCAATCAACGGAATGAGAGACTCTTGTCAACAGCAGACCTTGAGCCTC 417
Qy 46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValIleuTrp 65
Db 418 ACCGTCGTGACGTGATCTGTTCCCTTGTGGGCTGACAGAAACGAGTTGCTCTGG 477
Qy 66 LeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAla 85
Db 478 CTCCTGGGCTGCCACATGCGCAGAGAGCGCTTCTCAATCTCAATCTCAATCTGCGCA 537
Qy 86 AlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleAsnIle 105
Db 538 GCAAGCTTCTCTTCTCAAGCGCGCCCTTATATATCTCTGTTAAGCTTCATCAGATAC 597
Qy 106 ArgHisProIleSerIleLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
Db 598 CCCATACCATCTTAATCTTATCTCTGATGATGATGTTTCTTACCTTGACGGCTG 657
Qy 126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTrpProIleTyr 145
Db 658 AGCTTCTGAGTGGCGTAGACACGAGAGGCTGCTGCCGCTGCTGAGCCATCTGGTAC 717
Qy 146 HisGlyArgArgProArgTyrLeuSerSerValMetCysValIleLeuTrpAlaLeuSer 165
Db 718 CGCTGCCACCGCCACACACACTGTCAAGCGGTGTGTGTCTGCTGCGCCCTGTCC 777
Qy 166 LeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeuPheSerGlyAlaAspSer 185
Db 778 CTGCTGGGAGACATCTCGAGTGAATGTATGTGCTTCTCTCAATGAGTGAATTTCT 837
Qy 186 ValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuValPheLeuCysValIle 205
Db 838 GCTTGTGTCAAAATCAGATTTTCATCAGATCGCGGTGCTGATTTTATGTGTGTT 897
Qy 206 LeuCysGlySerSerIleuValIleuValArgIleLeuCysGlySerArgIleuMetPro 225
Db 898 CTCTGTGGGTTCACCGCTGTCTGTCTGATCAGATTTCTGTGTGATCCCGGAATATACG 957
Qy 226 LeuThrArgLeuTyrValIleThrIleLeuLeuThrValIleuValPheLeuLeuCysGlyLeu 245
Db 958 CTGACCAAGCTGTGACCATCTCTCACTACATCTGCTCTCTCTGTGGCTG 1017
Qy 246 ProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAspTrpIleValIleuPhe 265
Db 1018 CCCCTTGGCATTCAGTTTCTTATTTTATGATCCACGTCGACAGGGAAGTCTTATTT 1077
Qy 266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerAlaAsnProIle 285

```

```

Db 1078 TGTCAATTCATCTAGATTTTCTATTCTCTGCTTCAACAGCATGTCACCAACCCATC 1137
Qy 286 IleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgIleAsnLeuValIleuVal 305
Db 1138 ATTACTTCTTCTGCGGCTCTTATGACAGGTCAAAATATAGGCAACCTGATGGTT 1197
Qy 306 LeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGlyTyrTrpLeuProGln 325
Db 1198 CTCGAGGGCTCTGACGAGACGCTCTGAGTGAATAGTGAAGGAGGAGCTTCTGAG 1257
Qy 326 GluThrLeuGluLeuSerArgIleSerArgLeuGluGln 337
Db 1258 GAATCTCTGAGTCTGTCTGAGAGCAGATTTGAGAGC 1293

```

RESULT 14

US-10-017-161-1599
 Sequence 1599, Application US/10017161
 Publication No. US20030143668A1

GENERAL INFORMATION:
 APPLICANT: SUMA, MAKIKO
 APPLICANT: ASAI, KIYOSHI
 APPLICANT: AKIYAMA, YUTAKA
 APPLICANT: ABURATANI, HIROYUKI
 TITLE OR INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
 FILE REFERENCE: 084335/0152
 CURRENT FILING DATE: US/10/017.161
 PRIOR APPLICATION NUMBER: 2002-12-18
 PRIOR FILING DATE: 2001-06-18
 NUMBER OF SEQ ID NOS: 2430
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 1599
 LENGTH: 1370
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURES:
 NAME/KEY: source
 LOCATION: (1)..(1370)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (201)..(619)
 NAME/KEY: CDS
 LOCATION: (96)..(1170)
 US-10-017-161-1599

Alignment Scores:

Pred. No.:	7.3e-138	Length:	1370
Score:	1402.50	Matches:	293
Percent Similarity:	87.93%	Conservative:	13
Best Local Similarity:	84.20%	Mismatches:	29
Query Match:	79.55%	Indels:	15
	15	Gaps:	4

US-09-867-570-2 (1-337) x US-10-017-161-1599 (1-1370)

```

Qy 2 GluSerIleSerSerTrp-----ValIleArgLeuGly 12
Db 132 CAGAGATCAAAACGCTGTGATCATCATCTGTTCTGTTCAGAGGTCAATCAGCTGGG 191
Qy 13 PheLeuSerMetAspSerThrIleProValIleuGlyThrGluLeuThrProIleAsnGly 32
Db 192 TTTCTGAGCATGATTCACCATCTCCAGTCTTGGGTCAAGACTGACACCAATCAACGGA 251
Qy 33 ArgGluGluThrProCysTyrIleGlnThrLeuSerPheThrGlyLeuThrCysIleVal 52
Db 252 CGTAGGAGACTCTTGTCTAAGCAGACCTGAGCTTCAAGGGGCTGACGTGATGTT 311
Qy 53 SerLeuValAlaLeuThrGlyAsnAlaValIleuTrpLeuLeuGlyCysArgMetArg 72
Db 312 TCCCTTCTCGGCTGACAGAGAAAGCGGCTGTGTCTGCTCTCGGCTGCCGATCGCC 371

```

OY	73	ArganalaValSerIleTyrlleuamLeuValAlaAAspPheLeuPheuser	92
Db	372	AGGAACGCTGTCCATCATCATCTCAACCTGCTGGCGGCACTTCTCTCTTAC	431
OY	93	GlyHisIlelleCysSerProLeuArgLeuIleamIleArgHisProIleSerIle	112
Db	432	GGCCCAATTAATATGTGTGGCCGTAGCCCTCAATATCCGCCATCCATCCCAATTC	491
OY	113	LeuSerProValMetThrPheProTyrlPheIleGlyLeuSerMetLeuSerAlaIleSer	132
Db	432	CTCACTCTGTATATACCTTCTCCCTACTTATAGGCTTAAGCATCTGAGGCCATTCAGC	551
OY	133	ThrgluArgCysLeuSerIleLeuTrpProIleTrpTyrHisCysArgArgProArgTyr	152
Db	552	ACCGAGCGCTGCTGTCCATCTGTAGGCCCATCTGGTACATCTC-CGCGGCCAGATAC	610
OY	153	LeuSerSerValMetCysValLeuLeuTrpAlaLeuSerIleLeuArgSerIleLeuGlu	172
Db	611	CTGTATCGGTATATGTGTCTGTCTGTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT	670
OY	173	TrpMetPheCysAspPheLeuPheSerGly--AlaAspSerValTrpCysgluThrSer	191
Db	671	TGGATGATCTGTGACTACTCTGTATAGTGTGTCTGTATATGCTGGGTGTGTATAGCACACA	730
OY	192	AspPheIleThrIleAlaTrpLeuValPheLeuCysValValLeuCysGlySer--Ser	210
Db	731	GAC--ATTACAAACGCGGCTGTGTGTATATAGAGGGGTCTGCGGGGGGTGCAGC	787
OY	211	LeuValIleLeuLeuValArgIleLeuCysGlySerArgGlyMetProLeuThrArgLeu-Ty	230
Db	788	GTCTGTCTGTCTGTATAGATATGCGCGGATATGCGGAGATATGCCCTGATGATGTTCCAG	847
OY	230	rValThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleG1	250
Db	848	CGTGTATCATCTCTCCATGCTGTGTGTCTCTCTCCGCGGTGCTCTGCTCTTGG--ATTCA	905
OY	250	nTrpAlaLeuPheSerArgIleHisLeuAspTrpIlyValLeuPheCysHisValHisIle	270
Db	906	GTGGGCGCTGTITTCAGGATCCACTGGAGTGAAGTCTTATTTGTGCATGTACATCT	965
OY	270	uValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrlPhePheVa	290
Db	966	AGTTTCATTTTCTGTGCTCTTAAACAGAGTGCAACCCCATCATTTACTTCTTCT	1025
OY	290	IGlySerPheArgGlnArgGlnAsnArgGlnAsnIleuIleValLeuGlnArgAlaIle	310
Db	1026	GGGCTCTCTTTAGCGACGCTCAAAATAGGAGAACTGAAAGCTGTCTCCAGAGGGCTCT	1085
OY	310	uGlnAspThrProGluValAspGluGlyGlyTrpLeuProGlnIleuThrLeuGluIle	330
Db	1086	GCAGAGACACGCTGAGAGTGGATGAGAGGTGGAGGTGCTTCTCAGAGAACTCTGAGCT	1145
OY	330	uSerGlySerArgLeuGluGln	337
Db	1146	GTGCGGAGAGCAATGAGAGAG	1167
RESULT 15			
US-10-072-012-171			
/ Sequence 171, Application US/10072012			
/ Publication No. US20040033493A1			
GENERAL INFORMATION:			
/ APPLICANT: Tehernev, Velizar			
/ APPLICANT: Spytek, Kimberly			
/ APPLICANT: Zethusen, Bryan			
/ APPLICANT: Paturajan, Meera			
/ APPLICANT: Shimkete, Richard			
/ APPLICANT: Li, Li			
/ APPLICANT: Gangoli, Esha			
/ APPLICANT: Padigaru, Muralidhara			
/ APPLICANT: Anderson, David W.			
/ APPLICANT: Rastelli, Luca			
/ APPLICANT: Miller, Charles E.			
/ APPLICANT: Gerlach, Valerie			

```

1  APPLICANT: Taupier Jr, Raymond J.
2  APPLICANT: Gusev, Vladimir Y.
3  APPLICANT: Colman, Steven D.
4  APPLICANT: Wolenc, Adam R.
5  APPLICANT: Pena, Carol E. A
6  APPLICANT: Furtak, Katarzyna
7  APPLICANT: Grose, William M.
8  APPLICANT: Alsobrook II, John P.
9  APPLICANT: Lepley, Denise M.
10 APPLICANT: Rieger, Daniel K.
11 APPLICANT: Burgess, Catherine E.
12 TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
13 FILE REFERENCE: 21402-258
14 CURRENT APPLICATION NUMBER: US/10/072,012
15 CURRENT FILING DATE: 2002-01-31
16 PRIOR APPLICATION NUMBER: 60/265,102
17 PRIOR FILING DATE: 2001-01-30
18 PRIOR APPLICATION NUMBER: 60/265,514
19 PRIOR FILING DATE: 2001-01-31
20 PRIOR APPLICATION NUMBER: 60/265,517
21 PRIOR FILING DATE: 2001-01-31
22 PRIOR APPLICATION NUMBER: 60/265,412
23 PRIOR FILING DATE: 2001-01-31
24 PRIOR APPLICATION NUMBER: 60/265,395
25 PRIOR FILING DATE: 2001-01-31
26 PRIOR APPLICATION NUMBER: 60/266,406
27 PRIOR FILING DATE: 2001-02-02
28 PRIOR APPLICATION NUMBER: 60/266,767
29 PRIOR FILING DATE: 2001-02-05
30 PRIOR APPLICATION NUMBER: 60/267,057
31 PRIOR FILING DATE: 2001-02-07
32 PRIOR APPLICATION NUMBER: 60/266,975
33 PRIOR FILING DATE: 2001-02-07
34 PRIOR APPLICATION NUMBER: 60/267,459
35 PRIOR FILING DATE: 2001-02-08
36 Remaining SEQ ID Application data removed - See File Wrapper or PALM.
37 NUMBER OF SEQ ID NOS: 1391
38 SOFTWARE: PatentIn Ver. 2.1
39 SEQ ID NO 171
40 LENGTH: 997
41 TYPE: DNA
42 ORGANISM: Homo sapiens
43 US-10-072-012-171
44
45 Alignment Scores:
46 Pred. No.: 1,72e-137 Length: 597
47 Score: 1397.00 Matches: 272
48 Percent Similarity: 88.62% Conservative: 16
49 Best Local Similarity: 83.69% Mismatches: 37
50 Query Match: 79.24% Indels: 0
51 DB: 16 Gaps: 0
52
53 US-09-867-570-2 (1-337) x US-10-072-012-171 (1-997)
54
55 Oy 13 PheuserMeretAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGly 32
56 3 TTCTGTGAGCATGATTCACCAACATCTCAACTTGACACACAGAACTGACACCAATCAAGCA 62
57
58 Oy 33 ArgGluGluThrProCysTyrIysGlnThrLeuSerPheThrGlyLeuThrCysIleVal 52
59 63 ACTGAGGAGACTCTTTGCTACTCAAGCAGCACTTGAGCCTCAAGGCTGCTGACGCTGATCGT 122
60
61 Oy 53 SerLeuValAlaLeuThrGlyAsnAlaValIleLeuThrLeuLeuGlyCysArgMetArg 72
62 123 TCCCTTGTCCGCTGACAGGAACGACATGTGCTCTGCTCTCTCGGGCTGCGCATGCC 182
63
64 Db 73 ArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaIleAspPheLeuPheLeuSer 92
65 183 AGGAAGCGCTTCTTCATCTACATCTCAACTTGCGCGCAGACAGACTTCTCTCTCAGC 242
66
67 Oy 93 GlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHisproIleSerValIle 112
68 243 GAGCGCTTAAATTAATTCCTCGTAAGCTTCATCAAGATATCCCAACATCTTAATTC 302

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 1, 2004, 21:59:08 ; Search time 3989 Seconds

(without alignments)

3995.144 Million cell updates/sec

Title: US-09-867-570-2

Perfect score: 1763

Sequence: 1 MESKSSVIVLGFSLHMDST.....EGGGMPLQRTLELGSRLQ 337

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ .p2n.model -DEV=xlp
-Q=/cgm2.1/USPTO.spool.p/US09867570/runat.01122004.161122.6337/app.query.fasta_1.519
-DB=GenBank -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LCOPEXT=0 -LCOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USSR=US09867570.OCGN.1.1.5600 @runat.01122004.161122.6337 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenBank:
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1763	100.0	2618	6	AX429465 Sequence
2	1727.5	98.0	8622	6	AX429467 Sequence
3	1722.5	97.7	1369	6	AX647081 Sequence
4	1722.5	97.7	91923	9	AC108007 Homo sapi

C	5	1722.5	97.7	172939	9	AC090099	Homo sapi
	6	1721	97.6	1400	6	AX299158	Sequence
	7	1721	97.6	1400	6	AX549388	Sequence
	8	1721	97.6	1400	6	AY042215	Homo sapi
	9	1700	96.4	1697	9	BC067292	Homo sapi
	10	1688	95.7	969	6	E43450	Novel prote
	11	1688	95.7	969	6	E43450	Novel prote
	12	1683	95.5	969	6	E43451	Novel prote
	13	1683	95.5	969	6	AX923125	Sequence
	14	1675	95.0	969	6	AF474987	Homo sapi
	15	1642	93.1	969	6	AR477199	Sequence
	16	1642	93.1	969	6	BD103571	Novel G P
	17	1606	91.1	969	6	AF474988	Homo sapi
	18	1598	90.6	969	6	AR477200	Sequence
	19	1598	90.6	969	6	BD103572	Novel G P
	20	1416	80.3	1369	6	AX646705	Sequence
	21	1416	80.3	1369	6	AB065846	Homo sapi
	22	1416	80.3	2040	6	CO815414	Sequence
	23	1416	80.3	2040	6	AX299143	Sequence
	24	1416	80.3	2040	9	AY042213	Homo sapi
	25	1416	80.3	156839	9	AC107948	Homo sapi
	26	1416	80.3	163718	9	AC023078	Homo sapi
	27	1416	80.3	163718	9	AC139482	Homo sapi
	28	1411	80.0	969	9	AF474989	Homo sapi
	29	1405	79.7	969	6	AR477201	Sequence
	30	1405	79.7	969	6	BD103573	Novel G P
	31	1402.5	79.6	192618	2	AC020568	Homo sapi
	32	1399	79.4	912	6	AX657514	Sequence
	33	1397	79.2	997	6	AX921831	Sequence
	34	1390	78.8	1369	6	AX646849	Sequence
	35	1388	78.7	1604	6	AX299160	Sequence
	36	1388	78.7	1604	6	AX549403	Sequence
	37	1388	78.7	1604	9	AY042216	Homo sapi
	38	1387	78.4	969	6	E40003	Sequence
	39	1383	78.4	969	6	AX148178	Sequence
	40	1383	78.4	969	6	AX268489	Sequence
	41	1383	78.4	969	6	AX375428	Sequence
	42	1383	78.4	969	6	AX671113	Sequence
	43	1383	78.4	969	6	BD144297	Novel G-P
	44	1383	78.4	969	6	AB083628	Homo sapi
	45	1383	78.4	1116	6	BD095749	Novel gua

ALIGNMENTS

RESULT 1	AX429465	2618 bp	DNA	linear	PAT 21-UN-2002
LOCUS	AX429465	Sequence 1 from Patent WO0234914.			
DEFINITION	AX429465				
ACCESSION	AX429465.1	GI:21540763			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
TITLE	1				
JOURNAL	Zhao, Q., Beasley, E.M. and Wei, M.H.				
FEATURES	Isolated human g-protein coupled receptors, nucleic acid molecules				
source	encoding human gpcr proteins, and uses thereof				
	Patent: WO 0234914-A 1 02-MAY-2002;				
	PE CORP NY (US)				
	Location/Qualifiers				
	1..2618				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				

ORIGIN

Alignment Scores:

Pred. No.: 6.52e-137 Length: 2618

Score: 1763.00 Matches: 337

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-867-570-2 (1-337) x AX429465 (1-2618)

QY 1 MetGluSerLysSerSerTrpValIleArgLeuGlyPheLeuSerMetAspSerThrIle 20
 Db 447 ATGGAAAGCAATCTCTCATGGGTCATCAAGCTGGGTTTCTGAGCATTCATCAACCATC 506
 QY 21 ProValLeuGlyThrGluLeuThrProIleAsnGlyArgGluThrProCysTyrLys 40
 Db 507 CCAGCTTGCGGTACAGAACTGACCAATCAACGAGAGGAGAGACTCCTTCTCAAG 566
 QY 41 GlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsn 60
 Db 567 CAGACCTCGAGCTTCAACGGGGCTGAGAGTGCATGTTCCCTTTCGCGAGAGAGAAC 626
 QY 61 AlaValIleLeuThrLeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIle 80
 Db 627 GCGGTTGCTCTGGCTCTCTGGGCTCCGATCGAGAGAACCTGTCTTCATCTACATC 666
 QY 81 LeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyIleIleCysSerProLeu 100
 Db 687 CTCAACCTGCTCGCGGCGACCTTCTCTTCTTACGGGCGACATATATGTCGCCGCTTA 746
 QY 101 ArgLeuIleAsnIleArgHisProIleSerLysIleLeuSerProValMetThrPhePro 120
 Db 747 CGCCCTCATCAATATCGGCATCCCATCTCCAAATCTCAGTCTGTGATGACCTTTCC 806
 QY 121 TyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeu 140
 Db 807 TACTTATAGGCTTAAGCATGCTGAGCGCATACAGACGAGGCTGCCCTGCTCATCTG 866
 QY 141 TrpProIleTyrTyrHisCysArgArgProArgTyrLeuSerSerValMetCysValLeu 160
 Db 867 TGGCCCATCTGGTACACATGCGCGCCCGACAGATACCTGTCATCGATGATGTGCTG 926
 QY 161 LeuThrAlaLeuSerLeuLeuArgSerIleLeuGluThrMetPheCysAspPheLeuPhe 180
 Db 927 CTCTGGGCGCTGCTCTGCTGCGAGATCTGGAAGTGAATGTCGTGATCTTCGCTT 986
 QY 181 SerGlyAlaAspSerValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuVal 200
 Db 987 AGTGCTGCGATCTCTGTTGGTGTGAACCTCAATTCATTCATTCAGTGGTGGCTG 1046
 QY 201 PheLeuCysValIleLeuCysGlySerSerLeuValLeuValArgIleLeuCysGly 220
 Db 1047 TTTTATATGCTGCTCTGCTGAGGTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1106
 QY 221 SerArgLysMetProLeuThrArgLeuTyrValThrIleLeuLeuThrValIleValPhe 240
 Db 1107 TCCCGAAGATCCCGGTGACAGGCTGTAGTACCATCTCTCCACAGGCTGGCTTC 1166
 QY 241 LeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAsp 260
 Db 1167 CTCCTCTGGGCTGCTGCTTGGCATTCAGTGGGCTGTTTCCAGATCCACTGGAT 1226
 QY 261 TrpLysValIleLeuPheCysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSer 280
 Db 1227 TGGAAAGTCTTATTTTGTATGTCAGTACATAGTTTCCATTTTCTGTCGCTCTTAACAG 1286
 QY 281 SerAlaAsnProIleIleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgGln 300
 Db 1287 AGTGCAACCCCATATTAATCTTCGTGGGCTCTCTTAAGCAGGCTCAAAATAGAGCAG 1346
 QY 301 AsnLeuLysLeuValLeuGlnArgAlaLeuGlnAspThrProGluValAspGluGly 320
 Db 1347 AACCTGAAGCTGGTTCTCCAGAGGCTCTCCAGAGACAGCTTACAGTGAAGTGAAG 1406
 QY 321 GlyTyrPheProGlnGluThrLeuGluLeuSerGlySerArgLeuGluGln 337
 Db 1407 GGGTGGCTTCTTCAGGAAACCTCGAGCTGTGCGGAAGCAGATTGAGAGCAG 1457

RESULT 2
 AX429467
 LOCUS AX429467 8622 bp DNA linear PAT 21-JUN-2002

DEFINITION Sequence 3 from Patent WO0234914.
 ACCESSION AX429467
 VERSION AX429467.1 GI:21540764

KEYWORDS
 SOURCE
 ORGANISM

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Zhao, Q., Beasley, E. M. and Wei, M. H.

Isolated human g-protein coupled receptors, nucleic acid molecules

encoding human gpcr proteins, and uses thereof

Patent: WO 0234914-A 3 02-MAY-2002;

PE CORP NY (US)

FEATURES
 source 1. 8622
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 2,366-133 Length: 8622
 Score: 1727.50 Matches: 332
 Percent Similarity: 96.81% Conservative: 2
 Best Local Similarity: 96.23% Mismatches: 2
 Query Match: 97.99% Indels: 9
 DB: 6 Gaps: 1

US-09-867-570-2 (1-337) x AX429467 (1-8622)

QY 2 GluSerLysSerSerTrp-----ValIleArgLeuGly 12

Db 7459 CAGAGATCAACAGCGTGGATCAATCGTTCTGTTTCCAGGCTATCAAGCTGGGG 7518

QY 13 PheLeuSerMetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGly 32

Db 7519 TTTCTGAGCATGATTCACATCCAGCTTGGGTACAGAACTGACCAATCAACGGA 7578

QY 33 ArgGluGluThrProCysTyrLysGlnThrLeuSerPheThrGlyLeuThrCysIleVal 52

Db 7579 CGTAGGAGACTCTCTGCTACAAAGCAGACCTGAGCTTCAAGGCTGACGTCGCTT 7638

QY 53 SerLeuValAlaLeuThrGlyAsnAlaValLeuThrLeuLeuGlyCysArgMetArg 72

Db 7639 TCCCTGTGCGGCTGACAGAAACGGGTGCTCTGGCTCTGGGCTGGCGCATGGCG 7698

QY 73 ArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuSer 92

Db 7699 AGGAAGCTGTCTCCATCTCAATCTCAACCTGCTGCGGCGGACCTTCTCTCTTAC 7758

QY 93 GlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerLysIle 112

Db 7759 GGCACATTATATGTTGCGGCTTACGCTCATCAATATCCGACATCCATCTCCAAATC 7818

QY 113 LeuSerProValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSer 132

Db 7819 CTCAGTCTGTATGATGACTTCTTCTTACTTATAGGCTTAAGCATCTCAAGCGCATCAGC 7878

QY 133 ThrGluArgCysLeuSerIleLeuThrProIleTyrHisCysArgArgProArgTyr 152

Db 7879 ACCGAGCGTGTCTGCTCAATCTGTCGCTCACTGTGCTGCTGCTGCTGCTGCTG 7938

QY 153 LeuSerSerValMetCysValLeuLeuThrProIleAsnLeuSerLeuArgSerIleLeuGlu 172

Db 7939 CTGTATCGGCTATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7998

QY 173 TrpMetPheCysAspPheLeuPheSerGlyAlaAspSerValTrpCysGluThrSerAsp 192

Db 7999 TGGATGTTCTGACATTCCTGTTAGTGTGCTGATTCCTGTTGGTGTGAACCTGAGAT 8058
 Qy 193 PheIleThrIleAlaTrpLeuValPheLeuCyseValleuCyseGlySerSerLeuVal 212
 Db 8059 TTCATTACATCCGCTGGCTGGTGTGTTTATGATGTTCTGTGTGGTCCAGCTGTGTC 8118
 Qy 213 LeuLeuValArgIleLeuCyseGlySerArgIleMetProLeuThrArgLeuIleValThr 232
 Db 8119 CTGCTGTGAGATTCCTGTGTGATCCCGAAGATGCCGTGACAGCGTGTACGTACC 8178
 Qy 233 IleLeuLeuThrValleuValPheLeuLeuCyseGlyLeuProPheGlyIleGIntPAla 252
 Db 8179 ATCTCTCTCAAGATGCTGCTGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8238
 Qy 253 LeuPheSerArgIleIleIleLeuAspTrpValleuPheCysHisValHisLeuValSer 272
 Db 8239 CTGTTTCCAGATCCACCTGATGATGAAAGTCTTATTTGTCAATGATCATGATTTC 8298
 Qy 273 IlePheLeuSerIleLeuAsnSerSerAlaAsnProIleIleTyrrPhePheValGlySer 292
 Db 8299 ATTTTCTGCTGCTCTTAAACAGATGCCAACCCCACTTAACTTCTTCTGCTGCTGCTG 8358
 Qy 293 PheArgIleArgIleAsnArgIleAsnLeuLeuValleuGlnArgIleAlaLeuGlnAsp 312
 Db 8359 TTAGGCAAGCTCAAAATAGGCAAGCACTGAGCTGTTCTCCAGAGGCTCTGCAAGAC 8418
 Qy 313 ThrProGluValAspGluGlyIleGlyIleTrpLeuProGluGlnIleuThrLeuGluLeuSerGly 332
 Db 8419 ACCCTGAGGTGATGAAGGTGAGGGTGTGCTTCTCAGAAACCTGAGAGCTGTCCGGA 8478
 Qy 333 SerArgLeuGluGln 337
 Db 8479 AGCAGATTGAGCAG 8493

RESULT 3
 LOCUS AX647081 1369 bp DNA linear PAT 04-MAR-2003
 DEFINITION Sequence 1273 from Patent EP1270724.
 ACCESSION AX647081
 VERSION AX647081.1 GI:28800068

SOURCE
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 AUTHORS Suwa, M., Asai, K., Akiyama, Y. and Aburatani, H.
 TITLE Guanine ribophosphate-binding protein coupled receptors
 JOURNAL Patent: EP 1270724-A 1273 02-JAN-2003;
 National Institute of Advanced Industrial Science and Technology
 (JP) ; Center for Advanced Science and Technology Incubation, Ltd.
 (JP)

FEATURES
 source location/Qualifiers

CDS
 1..1369
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 201..1169
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD69666.1"
 /db_xref="GI:28800069"
 /translation="MDSTIPVLGTELPINGREBTPCYKOTLSFTGLTCTIVSVALTG
 NAVVLMGLCRMRNAVSIYILNLVADPFLSGHICSPRLINIRHPISKLSIPVM
 TPFYFGLSMWLSISTERCGLMPYICRPRPYLSWVCVLMALSLRSLIEMWF
 CDPLFGANSVWCSTSPRITIAMVPLFCVLCSSLVLYRIICGSRKMLPTLYVTI
 LLLVPLVLCGLPFGIOMAFSRHIDMKVLPCHVHVSIFLSANSSANPITTFYFG
 SFRQRQRQLKVLQALQDTPVDEGGMLPQETILSGSRLEQ"

ORIGIN
 Alignment Scores:
 Prod. No.:
 Score:

7.09e-134 Length: 1369
 1722.50 Matches: 331

Percent Similarity: 96.81% Conservative: 3
 Best Local Similarity: 95.94% Mismatches: 2
 Query Match: 97.70% Indels: 9
 DB: 6 Gaps: 1
 US-09-867-570-2 (1-337) x AX647081 (1-1369)

Qy 2 GluSerIleSerTrp-----ValIleGluGly 12
 Db 132 CAGAGATCAAAAGCTGTGTATGATCATCTGTTCTGTTCCAGGGGTCAAGCTGGGG 191
 Qy 13 PheLeuSerMetAspSerThrIleProValleuGlyThrGluLeuThrProIleAsnGly 32
 Db 192 TTTCTGAGCATGAGATTCACATCCAGTGTGGATGACACCTGACCAATCAAGGA 251
 Qy 33 ArgIleGluThrProCysTrpIleGlnThrLeuSerPheThrGlyLeuThrCysIleVal 52
 Db 252 CGTAGGAGACTCTTGCTTACAGAGACCCTGAGCTTCAAGGGGCTGACGTGATGTT 311
 Qy 53 SerLeuValAlaLeuThrGlyAsnAlaValleuTrpLeuGluGlyCysArgMetArg 72
 Db 312 TCCCTGTGCGCGCTGACAGGAAAGCGGTTGTCTCTGCTGCTGCTGCTGCTGCTGCTG 371
 Qy 73 ArgAsnAlaValSerIleTyrrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSer 92
 Db 372 AGAAGCGTGTCTCATCTACATCTCACCTGATGAGCGGCGGCACTTCTTCTTAC 431
 Qy 93 GlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIle 112
 Db 432 GGCACATTAATATGTTGCGGTTACGCTCATCAATATCGGCATCCATCTCCAAATC 491
 Qy 113 LeuSerProValMetThrPheProTyrrPheIleGlyLeuSerMetLeuSerAlaIleSer 132
 Db 492 CTCAGTCTGTGATGACCTTCCCTTATAGGCTTAAGCATGCTGACGCGCATCAGC 551
 Qy 133 ThrGluArgCysLeuSerIleLeuTrpProIleTrpTyrrHisCysArgGlyProArgTyrr 152
 Db 552 ACCGAGGCTGCTGCTCATCTGATCTGAGCCATGTGACACAGCGGCCGCCCAAGATAC 611
 Qy 153 LeuSerSerValMetCysValleuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGlu 172
 Db 612 CTGTCATCATGATCATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 671
 Qy 173 TrpMetPheCysAspPheLeuPheSerGlyAlaAspSerValTrpCysGluThrSerAsp 192
 Db 672 TGGATGTTCTGTGATCTTCTGTTAGTGTGATTAATCTGTTGGTGTGAACCTGAT 731
 Qy 193 PheIleThrIleAlaTrpLeuValPheLeuCyseValleuCyseGlySerSerLeuVal 212
 Db 732 TTCATTACATCCGCTGGCTGGTGTGTTTATGATGTTCTGTGTGGTCCAGCTGTGTC 791
 Qy 213 LeuLeuValArgIleLeuCyseGlySerArgIleMetProLeuThrArgLeuTyrrValThr 232
 Db 792 CTGCTGTGATGATTCCTGTGTGATCCGGAAGATGCCGTGACAGGCTGTACGTACC 851
 Qy 233 IleLeuLeuThrValleuValPheLeuLeuCyseGlyLeuProPheGlyIleGIntPAla 252
 Db 852 ATCTCTCTCAAGATGCTGCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 911
 Qy 253 LeuPheSerArgIleHisLeuAspTrpValleuPheCysHisValHisLeuValSer 272
 Db 912 CTGTTTCCAGATCCACCTGAGATGAAAGTCTTATTTGTCAATGATCATGATTTC 971
 Qy 273 IlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrrPhePheValGlySer 292
 Db 972 ATTTCTGCTGCTCTTAAACAGATGCCAACCCATCATTTACTTCTGCTGCTGCTGCTG 1091
 Qy 293 PheArgGlnArgGlnAsnArgGlnAsnLeuLeuValleuGlnArgAlaLeuGlnAsp 312
 Db 1032 TTAGGCAAGCTCAAAATAGGCAAGACCTGAACTGCTTCTCCAGAGGGCTCTGCAAGAC 1091
 Qy 313 ThrProGluValAspGluGlyIleGlyIleTrpLeuProGluGlnIleuThrLeuGluLeuSerGly 332

Db	1092	ACGGCTGAGTGAGTGAAGTGAGGAGTGCTTCTCAGAGAAACCTGAGCTGCGGGA	1151
Oy	333	SerATgLeUgUgLn 337	
Db	1152	AGCAGATTGGAGCAG 1166	
RESULT 4			
AC108007/c			
LOCUS	AC108007	91923 bp	DNA linear PRI 03-SEP-2002
DEFINITION	Homo sapiens chromosome 11, clone CTD-3038L12, complete sequence.		
ACCESSION	AC108007		
VERSION	AC108007.5	GI:22655822	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Enxaiyola; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 91923)		
TITLE	Birren,B., Nusbaum,C. and Lander,B.		
JOURNAL	Homo sapiens chromosome 11, clone CTD-3038L12		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 91923)		
Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barna,N., Baetien,V., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeKrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gaidyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kelle,C., Lacroque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., Mcwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,T., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Polara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tefaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.			
TITLE	Direct Submission		
JOURNAL	Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	3 (bases 1 to 91923)		
AUTHORS	Birren,B., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barna,N., Baetien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gaidyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kelle,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tefaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	4 (bases 1 to 91923)		
AUTHORS	Birren,B., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barna,N., Baetien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gaidyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kelle,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tefaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.		

```

TITLE
JOURNAL
COMMENT
    Cook,A., Cooke,P., Deatellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
    Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
    Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
    Karon,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
    Kataras,A., Kellis,C., Landers,T., Levine,R., Lindblad-Toh,K.,
    Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
    McCarthy,M., Meldrum,J., Menais,L., Mihova,T., Mings,V.,
    Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Nobu,C., Norman,C.H.,
    O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
    Phunhkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
    Roman,J., Roy,A., Schauer,S., Schupbach,R., Seaman,S., Severy,P.,
    Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamaa,J.,
    Teejaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
    Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
    Zembek,L., Zimmer,A. and Zody,M.
    Direct Submission
    Submitted (03-SEP-2002) Whitehead Institute/MIT Center for Genome
    Research, 320 Charles Street, Cambridge, MA 02141, USA
    On Sep 3, 2002 this sequence version replaced gi:12474954.
    All repeats were identified using RepeatMasker:
    Smt, A.P.A. & Green, P. (1996-1997)
    http://ftp.genome.washington.edu/RM/RepeatMasker.html
    ----- Genome Center
    Center: Whitehead Institute/ MIT Center for Genome Research
    Center code: WIBR
    Web site: http://www-seq.wi.mit.edu
    Contact: sequence_submissions@genome.wi.mit.edu
    ----- Project Information
    Center project name: L24596
    Center clone name: 3038_L_12
    -----
    Only the first 91.9 kb of this clone are being submitted.
    The remainder overlaps accession number AC055860 [WICGR project
    L9785].
    FEATURES
    SOURCE
        1. 91923
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="11"
        /map="11"
        /clone="CTD-3038L12"
        /clone_1lb="CTD2 Human BAC"
        25. 70
        /rpt_family="L2"
        complement(251.. 899)
        /rpt_family="L2"
        900..1001
        /rpt_family="L1MD2"
        1082..1655
        /rpt_family="L1MD2"
        1671..1850
        /rpt_family="L1MD3"
        complement(1858.. 1917)
        /rpt_family="L2"
        1918.. 2273
        /rpt_family="THE1B"
        complement(2274.. 2327)
        /rpt_family="L2"
        4505.. 4713
        /rpt_family="L2"
        5536.. 5795
        /rpt_family="AluSc"
        5796.. 5837
        /rpt_family="TAAAn"
        5838.. 5875
        /rpt_family="AluSc"
        complement(7186.. 7526)
        /rpt_family="L1MB4"
        complement(10716.. 10825)
        /rpt_family="MIR"
        11369.. 11466
    
```


AC090099/c	LOCUS	AC090099	172939 bp	DNA	linear	PRI 01-OCT-2002
	DEFINITION	Homo sapiens chromosome 11, clone RP11-113b6, complete sequence.				
	ACCESSION	AC090099				
	VERSION	AC090099.14	GI:23396332			
	KEYWORDS	HTG.				
	SOURCE	Homo sapiens (human)				
	ORGANISM	Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	AUTHORS	1 Birren,B., Nusbaum,C. and Lander,E.				
	TITLE	(bases 1 to 172939)				
JOURNAL		Homo sapiens chromosome 11, clone RP11-113b6				
REFERENCE	AUTHORS	2 (bases 1 to 172939)				
		Unpublished				
		Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,				
		Barna,N., Baerlein,V., Boguslavsky,L., Bouckgalter,B., Brown,A.,				
		Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,				
		Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,				
		Dodge,S., Fato,S., Ferreira,P., Fitzhugh,M., Gage,D., Galagan,J.,				
		Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,				
		Hagos,B., Hestford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,				
		Jones,C., Karatas,A., Larocque,K., Lamasares,R., Landers,T.,				
		Lenoczky,J., Levine,R., Liu,G., Maclean,C., MacDonald,P.,				
		Marquis,N., Matthews,C., McCarthy,M., Mewan,P., McKernan,K.,				
		Mccheeters,R., Meldrim,J., Menus,L., Mohwa,T., Mlenga,V.,				
		Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,				
		O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T., Peterson,K.,				
		Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,				
		Rieback,W., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,				
		Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P.,				
		Sounges,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,				
		Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,				
		Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A.,				
		Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,				
		Zembek,L., Zimmer,A. and Zody,M.				
TITLE	Direct Submission					
JOURNAL		Submitted (14-FEB-2001) Whitehead Institute/MIT Center for Genome				
REFERENCE	AUTHORS	Research, 320 Charles Street, Cambridge, MA 02141, USA				
		3 (bases 1 to 172939)				
		Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,				
		Barna,N., Baerlein,V., Bloom,T., Boguslavsky,L., Bouckgalter,B.,				
		Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A.,				
		Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,				
		Fato,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,				
		Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,				
		Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,				
		Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,				
		Liu,G., Maclean,C., MacDonald,P., Major,J., Matthews,C.,				
		McCarthy,M., Meldrim,J., Menus,L., Mihova,T., Mlenga,V.,				
		Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,				
		O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,				
		Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,				
		Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,				
		Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,				
		Testaye,S., Theodore,J., Tophan,K., Travers,M., Vassiliev,H.,				
		Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,				
		Zembek,L., Zimmer,A. and Zody,M.				
TITLE	Direct Submission					
JOURNAL		Submitted (05-SEP-2002) Whitehead Institute/MIT Center for Genome				
REFERENCE	AUTHORS	Research, 320 Charles Street, Cambridge, MA 02141, USA				
		4 (bases 1 to 172939)				
		Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,				
		Barna,N., Baerlein,V., Bloom,T., Boguslavsky,L., Bouckgalter,B.,				
		Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collamore,A.,				
		Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,				
		Fato,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,				
		Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,				
		Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,				
		Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,				
		Liu,G., Maclean,C., MacDonald,P., Major,J., Matthews,C.,				
		McCarthy,M., Meldrim,J., Menus,L., Mihova,T., Mlenga,V.,				
		Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,				
		O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,				
		Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,				
		Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,				
		Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,				

```

TITLE
JOURNAL
COMMENT

FEATURES
Source
Location/Qualifiers
1..172938
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-113D6"
/clone_1b="RP11 Human Male BAC"
/complement(1..77)
/rpt_family="AluS"
/complement(154..680)
/rpt_family="Char11e1"
/complement(678..810)
/rpt_family="Char1e1"
/complement(879..1057)
/rpt_family="L1M4"
/complement(1395..1513)
/rpt_family="FLAM_A"
1870..1876
/note="<30 qual SNGI region"
1902..1910
/note="<30 qual SNGI region"
2672..2788
/rpt_family="5S"
/complement(4009..4170)
/rpt_family="MIR"
/complement(4210..4355)
/rpt_family="MER5B"
4356..4647
/rpt_family="AluSx"
/complement(6554..6658)
/rpt_family="LTR41"
6916..6985
/rpt_family="MLT7B1"
/complement(7175..7359)
/rpt_family="LTR41"
/complement(8308..8611)
/rpt_family="AluYb"
8622..9083
/rpt_family="MLT7B1"
/complement(9265..10668)
/rpt_family="L1PB3"
10699..10809
/rpt_family="MER34"
10782..11264
/rpt_family="LTR48B"
/complement(11344..11683)
/rpt_family="L2"
11839..12026

```


Db 113981 TTCTGAGCATGATTCAACCATCCCACTTTGGGTACAGAACTGACACCAATCAACGGA 113922

RESULT	6
LOCUS	AX299158
DEFINITION	Sequence 30 from Patent WO018355.
ACCESSION	AX299158
VERSION	AX299158.1 GI:17129065
KEYWORDS	.
SOURCE	Homo sapiens (human)
PAT	26-NOV-2001 linear

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE

1 Anderson, D.J., Dong, X., Zylka, M., Han, S.K. and Simon, M.
Pain signaling molecules
Patent: WO 0183555-A 30 08-NOV-2001;
JOURNAL CALIFORNIA INSTITUTE OF TECHNOLOGY (US)

FEATURES

Location/Qualifiers

CDS

1..1400
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
332..1330
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD12713.1"
/db_xref="GI:17129066"
/translation="MDSITPVLATLTPINGREBPCKQTSFTGLTCTISLVALTG
NAVAVMLGCRMRRAVSIYILNLVADLPFGSHICSLRLINIRHPSKISLSPM
TFPPIGSLMSAISITERSLIMPIWYHCRPRYSYCVLNLALSLRSLIEMF
CDPLFSGADSVWCTSDPITIAIVPLCVLGGSLVILVRLICGRKMLRLVYTI
LLTVVPLICGIPFGIOWALFRIHIDMKVLCRHVLSIFISALNSANPILITPVVG
SFRQRONKMLVLPALQDTPEDVGGMLPQETLESGSRLEQ"

ORIGIN

Alignment Scores:

Pred. No.: 9.7e-134 Length: 1400
Score: 1721.00 Matches: 330
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 97.62% Indels: 0
DB: 6 Gaps: 0

US-09-867-570-2 (1-337) x AX293158 (1-1400)

QY 6 SerTPValIleArgLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThr 25
DB 302 TCCAGGGTCACACAGCTGGGTTTCTGAGCATGATTCACCATCCAGCTTGAGTAC 361
QY 26 GluLeuThrProIleAsnGlyArgGluGluThrProCysGlyTyrGlnThrLeuSerPhe 45
DB 362 GAACGACACCAATCAACGACGAGTGAAGAGACTCTTGCTACAGAGAACCTGAGCTTC 421
QY 46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValLeuLeuThr 65
DB 422 ACGGGGCTGAGCTGATCTCTTCCTTCGCGCTACAGAGAACCGGTTGCTCTG 481
QY 66 LeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAla 85
DB 482 CTCTGGGGCTGCGCATGCGACGAGAACGCTGTCTCATCTACATCCTCAACCTGCTCGCG 541
QY 86 AlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleAsnIle 105
DB 542 GCCGACTTCCTCTTCCTTACCGGCCACATTAATGTTCCGCTTACGCTCATCAATATC 601
QY 106 ArgHisProIleSerIleLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
DB 602 CGCCATCCCATCTCCAAATCTCTCAGTCTGTGATACACCTTCCCTTAATAGCCCTA 661
QY 126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuThrProIleTyrPyr 145
DB 662 AGCATGCTGAGCGCATGACACCGACGCGCTGCTCATCTGAGCCCATCTGCTAG 721
QY 146 HisCysArgArgProArgTyrLeuSerSerValMetCysValLeuLeuThrAlaLeuSer 165
DB 722 CACTGCGCGCGCCAGATACCTGTCACTGATCATGTGCTGTCTGCGGCCCTGTCTCC 781
QY 166 LeuLeuArgSerIleLeuGluThrMetPheCysAspPheLeuPheSerGlyAlaAspSer 185
DB 782 CTGCTGCGGAGATCTCTGAGTGGATGTTCTGTGATCTTCTGTGTTAGTGCTGATTTCT 841
QY 186 ValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuValPheLeuCysValVal 205

DB 842 GTTGTGTGAACGACGACATTCATTCACATCGGCTGCTGTTTATGTTGCTGTT 901
QY 206 LeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgLeuSerPro 225
DB 902 CTCTGAGGATCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961
QY 226 LeuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeu 245
DB 962 CTGACAGGCTGTACGTACATCTCTCCACAGAGCTGCTGCTCTCTCTCTGTCGCTG 1021
QY 246 PropheGlyIleGlnThrAlaLeuPheSerArgIleHisLeuAspTrpIleValLeuPhe 265
DB 1022 CCTTTGGCATTCAGTGGGCGCTGTTTCCAGGATCCACCTGAGTTGGAAGTCTTATTT 1081
QY 266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAspProIle 285
DB 1082 TGTCAATGTCATCTATGTTTCCATTTTCTGCTGCTCTTACACGACAGTCCACCCATTC 1141
QY 286 IleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgGlnAsnLeuLeuVal 305
DB 1142 ATTACTTCTTGTGCTGCTGCTCTTACGACGCTCAAAATAGGACGAACTGAAGCTGTT 1201
QY 306 LeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGlyTyrLeuProGln 325
DB 1202 CTCAGAGGCTCTCTGACAGACAGCGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1261
QY 326 GluThrLeuGluLeuSerGlySerArgLeuGluGln 337
DB 1262 GAACCTCTGAGCTGTGCGAAGACAGATTTGAGCAG 1297
RESULT 7
AX549388 1400 bp DNA linear PAT 26-NOV-2002
LOCUS AX549388
DEFINITION Sequence 673 from Patent WO02061087.
ACCESSION AX549388
VERSION AX549388.1 GI:25814004
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 Burner, G.C., Roush, C.J. and Brown, J.P.
Antigenic peptides, such as for G protein-coupled receptors
(GPCRs), antibodies thereto, and systems for identifying such
antigenic peptides
Patent: WO 02061087-A 673 08-AUG-2002;
JOURNAL Lifespan Biosciences, Inc. (US)
FEATURES
Location/Qualifiers
1..1400
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 9.7e-134 Length: 1400
Score: 1721.00 Matches: 330
Percent Similarity: 99.40% Conservative: 0
Best Local Similarity: 99.40% Mismatches: 2
Query Match: 97.62% Indels: 0
DB: 6 Gaps: 0
US-09-867-570-2 (1-337) x AX549388 (1-1400)
QY 6 SerTPValIleArgLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThr 25
DB 302 TCCAGGGTCACACAGCTGGGTTTCTGAGCATGATTCACCATCCAGCTTGAGTAC 361
QY 26 GluLeuThrProIleAsnGlyArgGluGluThrProCysGlyTyrGlnThrLeuSerPhe 45
DB 362 GAACGACACCAATCAACGACGAGTGAAGAGACTCTTGCTACAGAGAACCTGAGCTTC 421

QY 46 ThGlyLeuThrCysAlaValSerLeuValAlaLeuThrGlyAlaValValLeuTyr 65
 DB 422 ACCGGGGCTGACGGGATCGTTTCTTGTGCGCTGACAGAAACGGGTTGTGCTTGG 481
 QY 66 LeuLeuGlyCYsArpMetArGaArGaAlaValSerLeuTyrLeuAlaValAla 85
 DB 482 CTCCTGGGCTGGCGATGCGAGAGACGCTGTCTCACTCACTCACTCACTCACTGCG 541
 QY 86 AlaAspPheLeuPheLeuSerGlyAlaValCysSerProLeuArgLeuIleAsnIle 105
 DB 542 GCCGACTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 601
 QY 106 ArgHisProIleSerLeuValLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
 DB 602 CGCCATCCCACTCCCAAAATCCCTGATGCTGATGCTGATGCTGATGCTGATGCTG 661
 QY 126 SerMetLeuSerAlaIleSerThrGlyArgCysLeuSerIleLeuTyrProIleTyr 145
 DB 662 AGCATGCTGAGCGCATCAGACCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
 QY 146 HisCysArgArgProArgTyrLeuSerSerValMetCysValLeuLeuTyrAlaLeuSer 165
 DB 722 CACTGCGCGCGCGCGCATGCTGATGCTGATGCTGATGCTGCTGCTGCTGCTGCTG 781
 QY 166 LeuLeuArgSerIleLeuGlyTyrMetPheCysPhePheLeuPheSerGlyAlaAspSer 185
 DB 782 CTCCTCGGAGTATCTCGGAGTATCTCGGAGTATCTCGGAGTATCTCGGAGTATCT 841
 QY 186 ValTyrCysGlyLeuThrSerArpPheIleThrIleAlaTyrLeuValPheLeuCysValVal 205
 DB 842 GTTTGGTGGAAACGTACAGATTCATTACATGCGGCTGTTTATGATGATGTT 901
 QY 206 LeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgIleMetPro 225
 DB 902 CTCCTGGGCTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961
 QY 226 LeuThrArgLeuTyrValIleThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeu 245
 DB 962 CTCACACGAGCTGATGACGATCTCTCTCAAGTGTGCTCTCTCTCTCTCTCTCTG 1021
 QY 246 ProPheGlyIleGlnTyrAlaLeuPheSerArgIleHisLeuAspTyrPheValLeuPhe 265
 DB 1022 CCTTTGGACATCAAGGGGCTGTTTCCAGGATCCACCTGGATTGGAAAGCTTATTT 1081
 QY 266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIle 285
 DB 1082 TGTCAATGTCATCTAGTTTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1141
 QY 286 IleTyrPhePheValGlySerPheArgIleArgIleAsnArgIleAsnLeuValLeuVal 305
 DB 1142 ATTACT 1201
 QY 306 LeuGlyArgAlaLeuGlnAspThrProGlyValAlaGlyIleGlyTyrPheLeuProGln 325
 DB 1202 CTCACAGGGCTCTGACGAGACGCTGAGGTGATGAAGGTGAGGTGAGGTGAGGTG 1261
 QY 326 GluThrLeuGlyLeuSerGlySerArgLeuGlyGln 337
 DB 1262 GAAACCTGGAGCTGTGCGGAGAGATTGGAGAG 1297

RESULT 8
 AY042215 1400 bp DNA linear PRI 14-SEP-2001
 LOCUS Homo sapiens G protein-coupled receptor (MRGX3) gene, complete cds.
 DEFINITION
 ACCESSION AY042215
 VERSION AY042215.1 GI:15546065
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1400)

AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCES
 AUTHORS
 TITLE
 JOURNAL

Dong, X., Han, S., Zylka, M.J., Simon, M.I. and Anderson, D.J.
 A diverse family of GPCRs expressed in specific subsets of
 nociceptive sensory neurons
 Cell 106 (5), 619-632 (2001)
 21435808
 11551509
 2 (bases 1 to 1400)
 Dong, X., Han, S., Zylka, M.J., Simon, M.I. and Anderson, D.J.
 Direct Substitution
 Submitted (19-JUN-2001) Division of Biology, 216-76, California
 Institute of Technology, 1201 E. California Blvd., Pasadena, CA
 91125, USA

FEATURES
 source
 gene
 mRNA
 CDS

Location/Qualifiers
 1..1400
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <332..>1300
 /gene="MRGX3"
 <332..>1300
 /gene="MRGX3"
 /product="G protein-coupled receptor"
 332..1300
 /gene="MRGX3"
 /note="Mas-related gene X3"
 /codon_start=1
 /product="G protein-coupled receptor"
 /protein_id="AK91806.1"
 /db_xref="GI:15546066"
 /translation="MDSTIPVLTETLPINGREPTCYKQTLSPFTGLTCLVALNG
 MNYVLLGCRMRNRNASTYIINLVADPLFSGHISGRLINIRHPSKLSPM
 TPEYFGLSLAISTERCSTLIPWYHQRKRRYSVNCVLMASLRLSEKMF
 CDFLFGADSWCETSDFTIAMVFLCVLCCSSLVLLVRLICGRKMLPLTYVI
 LTLVTLFLCGLPFGIOMALFSRIHDMVKLFCHVHVSIFSLANSNAPIYFVG
 SFRQNRQNRQLKVLQALQDPTREVDEGGWLPQETLELSRLQG"

ORIGIN

Alignment Scores:
 Pred. No.: 9.7e-134 Length: 1400
 Score: 1721.00 Matches: 330
 Percent Similarity: 99.40% Conservative: 0
 Best Local Similarity: 99.40% Mismatches: 2
 Query Match: 97.62% Indels: 0
 DB: Gaps: 0

US-09-867-570-2 (1-337) x AY042215 (1-1400)

QY 6 SerTyrValIleArgLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThr 25
 DB 302 TCCAGGGGTCCAGACTGGGGGTTTCTGAGCATGATTCACCACTCCAGTCTGGGTACA 361
 QY 26 GluLeuThrProIleAsnGlyArgGlyGlnTyrProCysTyrIleGlnThrLeuSerPhe 45
 DB 362 GAATGACACCAATCAACGAGACCTGAGAGACTCTTGTCTACAGACGACCTGAGCTTC 421
 QY 46 ThGlyLeuThrCysAlaValSerLeuValAlaLeuThrGlyAlaValValLeuTyr 65
 DB 422 ACCGGGGCTGACGGGATCGTTTCTTGTGCGCTGACAGAAACGGGTTGTGCTTGG 481
 QY 66 LeuLeuGlyCYsArpMetArGaArGaAlaValSerIleTyrIleLeuAlaValAla 85
 DB 482 CTCCTGGGCTGGCGATGCGAGAGACGCTGTCTCACTCACTCACTCACTGCGG 541
 QY 86 AlaAspPheLeuPheLeuSerGlyAlaValCysSerProLeuArgLeuIleAsnIle 105
 DB 542 GCCGACTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 601
 QY 106 ArgHisProIleSerLeuValLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
 DB 602 CGCCATCCCACTCCCAAAATCCCTGATGCTGATGCTGATGCTGATGCTGATGCTG 661
 QY 126 SerMetLeuSerAlaIleSerThrGlyArgCysLeuSerIleLeuTyrProIleTyr 145

```

Db      662 AGCATGCTGAGCGGCATCAGACACGAGCGCTGCTTCATCTCTGAGCCCATCTGCTAC 721
Qy      146 HICGSAAGAGPPOHAGTGTITLSEUSERSVALMTCYVALLEULEUTPALALEUSER 165
Db      722 CACTGCCGCCGCCGACGATACCTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTC 781
Qy      166 LEULEUAGSERITLLEUGLUTRPMETPHECYASPPHELEUPHESERGLYALAASER 185
Db      782 CTGCGCGGAGATCTCGAGATGATGATCTCTGATCTCTGATCTCTGATCTCTGATCTCT 841
Qy      186 VALTRPGYGLUTRHSERASPPHEITLHRIEALTRPLEUVALPHELEUCYVALVAL 205
Db      842 GTTGAGTGTGAACGTCAGATTTCATTACATTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 901
Qy      206 LEUCYSGLYSERSERLEUVALLEUVALARYLLEUCYGLYSERARGYMETPRO 225
Db      902 CTCTGTGGTTCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 961
Qy      226 LEUTRARGLEUTY-VALTRITLLEULEUTRVALLEUVALPHELEUCYSGLYLEU 245
Db      962 CTGACCAAGGCTGATGATGATCCTCTCCACAGTCTGCTCTCTCTCTCTCTCTCTCT 1021
Qy      246 PROPHGLYILEGINTTPALALEUPHESERAGLIEHLEUASPTPLYSVALLEUPH 265
Db      1022 CCTTTGGCATTCACTGGGCGCTGCTTTCAGGATCCACTGGATTGGAAGCTTATTT 1081
Qy      266 CYSHSVALHISLEUVALSERITLPHLEUSERIALALEUASNSERIALAANPOLLE 285
Db      1082 TGTCAATGTCATCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1141
Qy      286 ILETYPHEPHEVALIGLYSERPHEARGINARGINASARGINASLEUVAL 305
Db      1142 ATTACTCTTCTGCTGGGCTCTTTAGGCGCGTCAAAATAGCAGAACCTGAGTGT 1201
Qy      306 LEUGLARGALALEUGLNASPTHPROGLIUALASPLUGLYGLYGLYTRPLEUPROGL 325
Db      1202 CTCGAGAGGCTCTCGACGACACGCTGAGTGATGAAGGCTGCTGCTCTCTCT 1261
Qy      326 GLUTRLEUGLLEUSERGLYSERARGLEUGLUGL 337
Db      1262 GAACCTGAGAGCTGCGGGAAGCATTTGAGCAG 1297

RESULT 9
BC067292 1697 bp mRNA linear PRI 30-JUN-2004
LOCUS Homo sapiens G protein-coupled receptor MRGX3, mRNA (CDNA clone
ACCESSION MGC:75450 IMAGE:30383708), complete cds.
VERSION BC067292
KEYWORDS BC067292.1 GI:45501076
SOURCE MGC.
ORGANISM Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 1697)
            Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
            Klausner, R.D., Collins, F.S., Wagner, K.H., Schaefer, C.F., Bhat, N.K.,
            Hopkins, R.F., Jordan, H., Moore, T., Mak, S.I., Wang, D., Hsieh, F.,
            Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
            Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
            Scheer, T.E., Brownstein, M.J., Uedlin, T.B., Toshiyuki, S.,
            Carinici, P., Prange, C., Rabe, S.S., Loquellano, N.A., Peters, G.J.,
            Abramson, R.D., Mullaney, S.J., Bosak, S.A., McEwan, P.J.,
            McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
            Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W.,
            Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
            Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
            Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
            Bouffard, G.G., Blakesley, R.W., Touchman, D.W., Green, E.D.,
            Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
            Butterfield, Y.S., Krzywinski, M.I., Skalske, U., Smalins, D.E.,
            Schermer, A., Schein, J.E., Jones, S.J. and Marra, M.A.

```

TITLE
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED
12477932
REFERENCE
2 (bases 1 to 1697)
AUTHORS
Strausberg, R.
TITLE
Direct Submission
JOURNAL
Submitted (12-MAR-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
NIR-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Hiao, Kim MacDonald, Amara Maasson, Mike R. Mayo, Josh Moran, Ryan Morin, Telia Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabh, Parvaneh Saeedi, JR Santos, Angelique Schermer, Ursula Skalske, Duane Smalls, Jeff Stott, Miranda Tsai, George Yang, Jacque Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAX Plate: 141 Row: m Column: 20
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 16876452.

FEATURES

source
1..1697
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:75450 IMAGE:30383708"
/issue_type="Testis"
/clone_1fb="NIR MGC_180"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
1..1697
/gene="MRGX3"
/db_xref="locusID:117195"
/db_xref="MIM:607229"
362..1330
/gene="MRGX3"
/codon_start=1
/product="G protein-coupled receptor MRGX3"
/protein_id="AAH67292.1"
/db_xref="GI:45501077"
/db_xref="locusID:117195"
/db_xref="MIM:607229"
/translacion="MDSITPVLTGTELPINGREBTPCYKOTLSFTGLTCLIVSLALNG
NAVVMILGCGRRRAVAISITLNLVLAAPFLSGHITRSPRLINIRHPSKILSPWM
TPPTITGMSLAIETECISLWPTWTHCRPRITSSVWVILMAISLISLIMNF
CDPLFSGDSVWCETSDITLAWLFLCVCGSLVLVILGSRMPPLRLIVTII
LITVLPFLICGLPFGIOWALPFSRHLMDKVLFCVHVLVSLFSLANSANPILYFPVG
SFRQRONRNLKLVLRALODTPVEDEGGCLPQETELSGSRLEQ"

ORIGIN

Alignment Scores:
Score: 6.77e-132
Percent Similarity: 1700.00
Best Local Similarity: 99.39%
Query Match: 96.43%
DB: 9
Length: 1697
Matches: 328
Conservative: 0
Mismatch: 2
Indels: 0
Gaps: 0

US-09-867-570-2 (1-337) x BC067292 (1-1697)

```

QY      8 ValIleArgLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThrGluLeu 27
DB      338 GTATATAGACTGAGGAGTTCTGAGACAGATTCACCACTCCAGCTTGAGTACGAACTG 397
QY      28 ThrProIleAsnGlyArgGluThrProCysTyrIleGlnThrLeuSerPheThrGly 47
DB      398 ACACCAATCAACGACACTGAGAGACTCTTGTCTACAGAGACCCCTGACCTTCAAGGG 457
QY      48 LeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValLeuThrLeu 67
DB      458 CTGACGCGCATCGTTCTCTGTCGCGCTGACAGAAACCGGGTGTGCTCTGCTCTG 517
QY      68 GlyCysArgMetArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaAlaAsp 87
DB      518 GGCTGCGCAGATGCGCAGAGAACGCTGCTTCATCTCAACCTCGTGGCGGCGAC 577
QY      88 PheLeuPheLeuSerGlyHisIleCysSerProLeuArgLeuIleAsnIleArgHis 107
DB      578 TTCTCTTCTTCTTACGGCCCATTTATCGTTCCGCGTTACGCTCATCAATATCCGCAT 637
QY      108 ProIleSerIleValLeuSerProValMetThrPheProTyrPheIleGlyLeuSerMet 127
DB      638 CCCATCTCCAAATTCCTCAGTCCGTGATGACCTTCCCTACTTATAGCCTTAAGCATG 697
QY      128 LeuSerAlaIleSerThrGluArgCysLeuSerIleLeuThrProIleTyrIleCys 147
DB      698 CTGAGGCGCCATCGACACCGAGCGCTGCTCCATCTGAGCCCATCTGTTACCTGCG 757
QY      148 ArgArgProArgTyrLeuSerSerValMetCysValLeuLeuThrAlaLeuSerLeu 167
DB      758 CGCGCGCCGAGATACCTGATCGGTATGCTGCTCTGCGCGCCCTGCTCTCTG 817
QY      168 ArgSerIleLeuGlnTyrPheCysAspPheLeuPheSerGlyAlaAspSerValTyr 187
DB      818 CGAGATATCTGAGATGATGTTCTGTGATCTCTGTTAGTGTGATGATTCGTTTG 877
QY      188 CysGlnThrSerPheIleThrIleAlaTyrPheValPheLeuCysValValLeuCys 207
DB      878 TGTGAACGTCACATTTTCATTAACATCGCGTGTGTTTATGCTGTGCTCTCTG 937
QY      208 GlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgIleMetProLeuThr 227
DB      938 GGGTTCACGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 997
QY      228 ArgLeuTyrValThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeuProPhe 247
DB      998 AGGCTGAGCTGACCATCTCTCAAGTCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1057
QY      248 GlyIleGlnTyrAlaLeuPheSerArgIleHisLeuAspTyrPheValLeuPheCysHis 267
DB      1058 GGCAATTCAGAGGCGCCGTTTTCAGGATCCACCTGATGGAAGCTTATTTTGTGCAT 1117
QY      268 ValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyr 287
DB      1118 GTCACTATATTCATTTCTCTGCTGCTCTTAAACGACGTCGCAACCCCATCATTTTAC 1177
QY      288 PhePheValGlySerPheArgGlnArgGlnAsnArgGlnAsnLeuLeuValLeuGln 307
DB      1178 TTCTTGTGGGCTCTCTTATGACGCGTCAAAATAGAGAACTGAAAGCTGCTTCTCAG 1237
QY      308 ArgAlaLeuGlnAspThrProGluValAspGluGlyGlyTyrPheLeuProGlnGluThr 327
DB      1238 AGGGCTCTGAGAGACGCTGAGGTGATGAAGTGAAGGAGGAGCTTCTCTCAAGAAACC 1297
QY      328 LeuGluLeuSerGlySerArgLeuGluGln 337
DB      1298 CTGAGAGCTGTGGAGACAGATTGAGACAG 1327

```

```

RESULT 10
LOCUS   E43450          969 bp    DNA          linear    PAT 31-JAN-2002
DEFINITION Novel protein G-coupled receptor protein and DNA thereof.

```

```

ACCESSION E43450
VERSION   E43450.1 GI:18627716
KEYWORDS  JP 2000166576-A/1.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 969)
AUTHORS   Watanabe,T., Terao,S. and Matsui,H.
TITLES    Novel protein G-coupled receptor protein and DNA thereof
JOURNAL   Patent: JP 2000166576-A 1 20-JUN-2000;
          TAKEDA CHEM IND LTD
COMMENT    OS Homo sapiens (human)
           PN JP 2000166576-A/1
           PD 20-JUN-2000
           PR 30-SEP-1999 JP 1999278116
           PR TAKUYA WATANABE, SHIZUKO TERAOKA, HIDEKI MATSUI
           PC C12N15/09,A61K38/00,A61K39/395,A61K45/00,A61K48/00,
           PC A61P43/00,
           PC C07K14/705,C07K16/28,C12N1/21,C12N5/10,C12P21/02,C12P21/08,PC
           PC G01N33/15,
           PC G01N33/53,G01N33/566,G01N33/577/(C12N1/21,C12R1:19),(C12N5/10,PC
           C12R1:91),
           PC (C12P21/02,C12R1:19),(C12P21/08,C12R1:91),C12N15/00,A61K37/02,
           PC C12N5/00,
           PC (C12N5/00,C12R1:91)
CC         CC
FH         Key location/Qualifiers
FT         source 1..969
           /organism='Homo sapiens (human)'.
FEATURES   source
           location/Qualifiers
           1..969
           /organism='Homo sapiens'
           /mol_type='genomic DNA'
           /db_xref='taxon:9606'
ORIGIN
Alignment Scores:
Pred. No.: 3..49e-131 Length: 969
Score: 1688.00 Matches: 322
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.75% Indels: 0
DB: Gaps: 0
US-09-867-570-2 (1-337) x E43450 (1-969)
QY      16 MetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGlyArgGluGln 35
DB      1 ATGATTCACCAACATCCACAGCTTGGGTGACAGACTGACCAATTCACAGACCTGAGAG 60
QY      36 ThrProCysTyrIleArgGlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
DB      61 ACTCTTGTGACAGAGACGACCTGAGCTTCACGAGGCTGACGTCATCGTTCCCTTGTG 120
QY      56 AlaLeuThrGlyAsnAlaValValLeuThrLeuGlyCysArgMetArgAsnAla 75
DB      121 GCGCTGACAGAAACCGGCTGTGTGCTCTGCTCTGCGGCGCGCATGCCAGAAAGCT 180
QY      76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95
DB      181 GTCTCATCTACATCTCAACCTGCTGCGCGCGCATCTCTCTTACGGCCCATAT 240
QY      96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
DB      241 ATATGTTCCGCTTACGCTCATCAATATCCGCAATCCCAATCCCAATCTCAAGTCT 300
QY      116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
DB      301 GTATATACCTTCTCTTATAGGCTTAAGCATGTGAGCGGCATCAACGAGCGC 360

```

QY 136 CysLeuSerIleLeuTrpProIleTrpIleHisCysArgArgProArgIleLeuSerSer 155
 Db 361 TGCCTGTCATCCTTGGCCCATCTGTGATACCACTGCGCGGCCGACGAACTCTGATCG 420
 QY 156 ValMetCysValIleuLeuTrpAlaLeuSerIleuLeuArgSerIleuGluTrpMetPhe 175
 Db 421 GTCAATGTGTCTCTCTGCGGCCCTGTCTCCCTGTGCGAGATACCTGTGAGTGAATGTC 480
 QY 176 CysAspPheLeuPheSerGlyValIleAspSerValITrPcysGluThrSerAspPheIleThr 195
 Db 481 TGTAACTTCTGTGTAGTGTGCTGATATCTGTTGTGTGAACGTCAGATTTCATTACA 540
 QY 196 IleAlaTrpLeuValPheLeuCysValIleuCysGlySerSerLeuValIleuLeuVal 215
 Db 541 ATCCGCTGCTGCTTTTATGTGTGTCTCTGTGGGTCCAGCCTGCTGCTGCTGCTGCTG 600
 QY 216 ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTrpValThrIleLeuLeu 235
 Db 601 AGCAATCTCTGTGATCCCGGAAAGATGCGCTGACAGAGCTGTACGTGACCAATCCTCTC 660
 QY 236 ThrValIleuValPheLeuLeuCysGlyLeuProPheGlyIleGluTrpAlaLeuPheSer 255
 Db 661 ACAGAGTGTGCTCTCTCTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 QY 256 ArgIleHisIleuAspTrpIleValIleuPheCysHisValIleuValSerIlePheLeu 275
 Db 721 AGGATCCACTGTGATTTGAAAGTCTTATTTGTGATGTCATCTGATTTCCATTTCCG 780
 QY 276 SerAlaLeuAsnSerSerAlaAsnProIleIleTrpPhePheValGlySerPheArgIle 295
 Db 781 TCCGCTCTTAAACAGAGTGCACCAACCCCATCTTCTGCTGCTGCTGCTGCTGCTGCTG 840
 QY 296 ArgGlnAsnArgGlnAsnLeuLeuValIleuGlnArgAlaLeuGlnAsnTrpProGlu 315
 Db 841 GGTCAAAATAGGCAAGAACTGAAAGCTGTCTTCACAGAGCTGTGAGAGACAGCTGTAG 900
 QY 316 ValAspGluGlyGlyIleTrpLeuProGluGluThrLeuGluLeuSerGlySerArgLeu 335
 Db 901 GTGATGAAGGTGAGAGGTGCTTCTCTCAAGAAACCTGTGAGCTGTGCGGAAGCAATTTG 960
 QY 336 GluGln 337
 Db 961 GAGCAG 966
 RESULT 11
 LOCUS AX498196 969 bp DNA linear PAT 26-SEP-2002
 DEFINITION Sequence 19 from Patent WO0242461.
 ACCESSION AX498196
 VERSION AX498196.1 GI:23343119
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM artificial sequences.
 REFERENCE 1
 AUTHORS Chen, R., Chu, Z. L., Dang, H. T., Lowitz, K. P. and Pride, C.
 TITLE Endogenous and non-endogenous versions of human g protein-coupled
 JOURNAL Patent: WO 0242461-A 19 30-MAY-2002;
 Arena Pharmaceuticals, Inc. (US)
 FEATURES
 source 1..969
 Location/Qualifiers
 1..969
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="Novel Sequence"
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,49e-131 Length: 969
 Score: 1688.00 Matches: 322
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 95.75% Indels: 0
 DB: 6 Gaps: 0
 US-09-867-570-2 (1-337) x AX498196 (1-969)
 QY 16 MetAspSerThrIleProValIleuGlyThrGluLeuThrProIleAsnGlyArgGlu 35
 Db 1 ATGATATCAACCATCCCAAGTCTTGGGTACAGAACTGACCAATCAACGAGCAGAGAG 60
 QY 36 ThrProCysTrpIleGluThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
 Db 61 ACTCTGTGTACAAACAGACCTGAGCTTCAAGGAGCTGAGCTGATGCTGTTCCCTGTGC 120
 QY 56 AlaLeuThrGlyAsnAlaValIleuTrpLeuLeuGlyCysArgMetArgArgAsnAla 75
 Db 121 GCGTGACAGAAAGCGGTTGTCTGTGCTCTCGGCTCCCGATGCGAGAGAGCT 180
 QY 76 ValSerIleTrpIleLeuAsnLeuValAlaIleAspPheLeuPheLeuSerGlyHisIle 95
 Db 181 GTCTCCATCTCATCTCAACCTGAGTCCGCGGCCACTTCTCTTACGCGGCACACTT 240
 QY 96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerValIleuSerPro 115
 Db 241 ATATGTGCGCCGTACGCTCATCAATCCGCAATCCCATCTCCAAATCTCAATCTCT 300
 QY 116 ValMetThrPheProTrpPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
 Db 301 GTGATGACCTTTCCCTACTTTATAGGCTTATAGCATGTCAGACGCCATGACGACGAGCG 360
 QY 136 CysLeuSerIleLeuTrpProIleTrpIleHisCysArgArgProArgIleLeuSerSer 155
 Db 361 TGCCTGTCATCCTTGGCCCATCTGTGATACCACTGCGCGGCCGACGAACTCTGATCG 420
 QY 156 ValMetCysValIleuLeuTrpAlaLeuSerIleuLeuArgSerIleuGluTrpMetPhe 175
 Db 421 GTCAATGTGTCTCTCTCTGCGGCCCTGTCTCCCTGTGCGAGATACCTGTGAGTGAATGTC 480
 QY 176 CysAspPheLeuPheSerGlyValIleAspSerValITrPcysGluThrSerAspPheIleThr 195
 Db 481 TGTAACTTCTGTGTAGTGTGCTGATATCTGTTGTGTGAACGTCAGATTTCATTACA 540
 QY 196 IleAlaTrpLeuValPheLeuCysValIleuCysGlySerSerLeuValIleuLeuVal 215
 Db 541 ATCCGCTGCTGCTTTTATGTGTGTCTCTGTGGGTCCAGCCTGCTGCTGCTGCTGCTGCTG 600
 QY 216 ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTrpValThrIleLeuLeu 235
 Db 601 AGCAATCTCTGTGATCCCGGAAAGATGCGCTGACAGAGCTGTACGTGACCAATCCTCTC 660
 QY 236 ThrValIleuValPheLeuLeuCysGlyLeuProPheGlyIleGluTrpAlaLeuPheSer 255
 Db 661 ACAGAGTGTGCTCTCTCTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 QY 256 ArgIleHisIleuAspTrpIleValIleuPheCysHisValIleuValSerIlePheLeu 275
 Db 721 AGGATCCACTGTGATTTGAAAGTCTTATTTGTGATGTCATCTGATTTCCATTTCCG 780
 QY 276 SerAlaLeuAsnSerSerAlaAsnProIleIleTrpPhePheValGlySerPheArgIle 295
 Db 781 TCCGCTCTTAAACAGAGTGCACCAACCCCATCTTCTGCTGCTGCTGCTGCTGCTGCTG 840
 QY 296 ArgGlnAsnArgGlnAsnLeuLeuValIleuGlnArgAlaLeuGlnAsnTrpProGlu 315
 Db 841 GGTCAAAATAGGCAAGAACTGAAAGCTGTCTTCACAGAGCTGTGAGAGACAGCTGTAG 900
 QY 316 ValAspGluGlyGlyIleTrpLeuProGluGluThrLeuGluLeuSerGlySerArgLeu 335
 Db 901 GTGATGAAGGTGAGAGGTGCTTCTCTCAAGAAACCTGTGAGCTGTGCGGAAGCAATTTG 960
 QY 336 GluGln 337
 Db 961 GAGCAG 966

RESULT 12
E43451 969 bp DNA linear PAT 31-JAN-2002
LOCUS Novel protein G-coupled receptor protein and DNA thereof.
DEFINITION E43451.1 GI:18627717
ACCESSION E43451.1 GI:18627717
VERSION JP 2000166576-A/2.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Euteheria; Primates; Carnivora; Homnidae; Homo.
1 (bases 1 to 969)
REFERENCE Watanabe, T., Terao, S. and Matsui, H.
Novel protein G-coupled receptor protein and DNA thereof
Patent: JP 2000166576-A 2 20-JUN-2000;
TAKEDA CHEM IND LTD
COMMENT OS Homo sapiens (human)
PN JP 2000166576-A/2
PD 20-JUN-2000
PP 30-SEP-1999 JP 1999278116
PR
PI TAKUYA WATANABE, SHIZUKO TERAOKA, HIDEKI MATSUI
PC C12N15/09, A61K38/00, A61K39/395, A61K45/00, A61K48/00,
PC A61P43/00,
PC C07K14/705, C07K16/28, C12N1/21, C12N5/10, C12P21/02, C12P21/08, PC
G01N33/15,
PC G01N33/53, G01N33/566, G01N33/577// (C12N1/21, C12R1.19), (C12N5/10, PC
C12R1.91),
PC (C12P21/02, C12R1.19), (C12P21/08, C12R1.91), C12N5/00, A61K37/02,
PC C12N5/00,
PC (C12N5/00, C12R1.91)
CC
CC
CC
FT source location/Qualifiers
FT 1..969 location/Qualifiers
FT /organism='Homo sapiens (human)'.
FEATURES
source location/Qualifiers
1..969
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Alignment Scores:
Prid. No.: 9.08e-131 Length: 969
Score: 1683.00 Matches: 321
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.69% Mismatches: 0
Query Match: 95.46% Indels: 0
DB: Gaps: 0
US-09-867-570-2 (1-337) x E43451 (1-969)
QY 16 MetAspSerThrIleProValLeuGluThrGluLeuThrProIleAsnGluYArgGluGlu 35
DB 1 ATGGATTCAACCAATCCCAAGTCTTGAGTACAAATGACCAACAAACGAGCGAGAG 60
QY 36 ThrProCysArgTyrLysGlnThrLeuSerPheThrGlyLeuThrCysAlaValSerLeuVal 55
DB 61 ACTCCCTTGACAAAG 120
QY 56 AlaLeuThrGlyAlaValValLeuThrLeuGluGlyCysArgMetArgGlnAla 75
DB 121 GGCCTGACAGAAACCGGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuSerGlyIleIle 95
DB 181 GTCTCCATTCACCAATCCCAAGTCTTGAGTACAAATGACCAACAAACGAGCGACATT 240
QY 96 IleCysSerProLeuArgLeuIleAsnIleArgIleSerProIleSerIleLeuSerPro 115
DB 241 ATATGTTGGCGTTAGCGCTCATCAATATCGGCATCCCATCTCCAAATCTCAGTCT 300

QY 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
DB 301 GTGATGACCTTCCCACTTATAGGCTTAAGATGCTGAGCGGCATCAGACGACGAC 360
QY 136 CysLeuSerIleLeuTyrProIleTyrPheIleCysArgArgProArgTyrLeuSerSer 155
DB 361 TGCCTGCTCAATCTCTGAGCCCACTGATCAGTCCGCGCCCAATACCTTCATCG 420
QY 156 ValMetCysValLeuLeuTyrPalaLeuSerLeuLeuArgSerIleLeuGluTyrMetPhe 175
DB 421 GTCATGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThr 195
DB 481 TGTGACTTCCCTGTTTATGTGTGCTAAATCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 540
QY 196 IleAlaTyrLeuValPheLeuCysValIleLeuCysGlySerSerLeuValIleLeuVal 215
DB 541 ATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 216 ArgIleLeuCysGlySerArgTyrMetProLeuThrArgLeuTyrValIleLeuLeu 235
DB 601 AGGATTCCTGTGTGATCCCGAAGATGCGCTGACGAGCTGTGACATCATCTCTC 660
QY 236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTyrAlaLeuPheSer 255
DB 661 ACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 256 ArgIleIleLeuAspTyrLysValIlePheCysHisValIleLeuValSerIlePheLeu 275
DB 721 AGATTCACCTGATGATGAAAGCTTATTTGTGATGATGATGATGATGATGATGATGAT 780
QY 276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGln 295
DB 781 TCGGCTCTTAACAGAGATGCAACCCCATCATTTACTTTCGAGGCTCTTAAAGCAG 840
QY 296 ArgGlnAsnArgGlnAsnLeuValLeuValLeuGlnArgAlaLeuGlnAsnThrProGlu 315
DB 841 CGTCAAATATGAGCAAGACCTGAAAGCTGTCTCCAGAGGCTCTGACGAGACACCCGAG 900
QY 316 ValAspGluGlyGlyGlyTyrLeuProGluGluThrLeuGluLeuSerGlySerArgLeu 335
DB 901 GTGATGTAAGTGAAGGATGCTTCTCAGAAACCTGAGCTGTGGAAGACATTTG 960
QY 336 GluGln 337
DB 961 GAGCAG 966
RESULT 13
AX923125 969 bp DNA linear PAT 18-DEC-2003
LOCUS Sequence 1 from Patent WO03080659.
DEFINITION AX923125
ACCESSION AX923125
VERSION AX923125.1 GI:40216228
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Euteheria; Primates; Carnivora; Homnidae; Homo.
1 (bases 1 to 969)
REFERENCE Peri, K.G., Moffett, S. and Abbran, D.
Methods and compounds for prevention and treatment of elevated
intraocular pressure and related conditions
Patent: WO 03080659-A 1 02-OCT-2003;
THERATECHNOLOGIES INC. (CA)
COMMENT OS Homo sapiens (human)
PN JP 2000166576-A/2
PD 20-JUN-2000
PP 30-SEP-1999 JP 1999278116
PR
PI TAKUYA WATANABE, SHIZUKO TERAOKA, HIDEKI MATSUI
PC C12N15/09, A61K38/00, A61K39/395, A61K45/00, A61K48/00,
PC A61P43/00,
PC C07K14/705, C07K16/28, C12N1/21, C12N5/10, C12P21/02, C12P21/08, PC
G01N33/15,
PC G01N33/53, G01N33/566, G01N33/577// (C12N1/21, C12R1.19), (C12N5/10, PC
C12R1.91),
PC (C12P21/02, C12R1.19), (C12P21/08, C12R1.91), C12N5/00, A61K37/02,
PC C12N5/00,
PC (C12N5/00, C12R1.91)
CC
CC
CC
FT source location/Qualifiers
FT 1..969 location/Qualifiers
FT /organism='Homo sapiens'
/mol_type='unassigned DNA'
/db_xref='taxon:9606'
/note='unnamed protein product'
/codon_start=1


```

/protein_id="CAF02233.1"
/db_xref="GI:40216229"
/translation="MDSTIPVGLTELTPINGREPTCYKQTLSPFGLTIVSLVALTG
DAVVMILGCRMRNAVSIYIINLVAAPFLFSGHIICSPILINIRHISIKLSPVM
TPPFIIGLSMLSAISTERCLSLMPDIWYHCRPRRLYSVCLVLAIRLSILSPM
CDFLPGSGNSWCERSDPTITAMLVFLCVLCSGLVLRVILCSRKMLPRLVYVI
LITVAVFLICGLPFGIOWALFSRIHDMKVLFCVHALVIFPSALNSANPITTFPVG
SFRORONRNLKIVLRALQDTPVDEGGMLPQETLISGSRLEQ"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 9.08e-131      Length: 969
Score: 1683.00      Matches: 321
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.69%      Mismatches: 0
Query Match: 95.46%      Indels: 0
DB: 6      Gaps: 0

```

US-09-867-570-2 (1-337) x AX923125 (1-969)

```

QY 16 MetAapSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGlyArgGluGlu 35
DB 1 AVGATTTCAACCATCCAGTCTTGAGTACAGAACGACCAATCAACGACGAGAGAG 60
QY 36 ThrProCysTyrGlyGlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
DB 61 ACTCTTCTCTACACAGACACCTGAGCTTCAAGGGCTGACGTGCATCTTCTTCTGTC 120
QY 56 AlaLeuThrGlyAsnAlaValLeuThrLeuGlyCysArgMetArgArgAsnAla 75
DB 121 GCGCTGACAGAGACGCGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95
DB 181 GCTTCATCTACATCTCAACTGAGTGGCGGCGGACCTTCTTCTTGAAGCGGCACAT 240
QY 96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
DB 241 ATATGTTGCGCGTTAGCGCTCATCATATCCGCAATCCCAATCCCAATCCCAATCCCT 300
QY 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
DB 301 GTGATGACCTTCCCTACTTATAGGCTTAAGGCTTAAGGCTTAAGGCTTAAGGCTTAAG 360
QY 136 CysLeuSerIleLeuThrProIleTyrPheHisCysArgArgProArgTyrLeuSerSer 155
DB 361 TGCCTGTCATCTGAGGCCCATCTGAGTACCACTGCGCGCGCCGACGATACCTGTATCG 420
QY 156 ValMetCysValLeuLeuThrPheIleLeuSerLeuLeuArgSerIleLeuGluThrMetPhe 175
DB 421 GTCAATGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThr 195
DB 481 TGTGATCTTCCTGTTAGTGTGCTGATTCGTTTGTGTAACACGACGATTTTCATTAC 540
QY 196 IleAlaThrLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuVal 215
DB 541 ATCCGCTGAGCTGCTTTTATGTTGTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 600
QY 216 ArgIleLeuCysGlySerArgLeuMetProLeuThrArgLeuTyrValThrIleLeuLeu 235
DB 601 AGGATTTCTCTGATGCCGGAAGATGCCGCTGACCAAGGCTGATGACCAATCTCTCT 660
QY 236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyTyrIleGlnThrAlaLeuPheSer 255
DB 661 ACAGTGTGCTGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 256 ArgIleHisLeuAspThrPheValLeuPheCysHisValHisLeuValSerIlePheLeu 275
DB 721 AGGATTCACCTGATGTAAGCTTAATTTGTCATGTGATCTAGATTTTCATTCTCTG 780
QY 276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGln 295

```

```

DB 781 TTCGCTTTAAGACAGAGGCCAACCCCATCTTACTTCTGAGGCTCTCTTAGGACAG 840
QY 236 ArgGlnAsnArgGlnAsnLeuLeuValLeuGlnAlaGlnAlaGlnThrProGlu 315
DB 841 CTTCAAAATAGGACAACTGAACTGTTCTTCCAGAGGCTCTGACAGACACCTTAG 900
QY 316 ValAspGluGlyGlyTyrPheProGlnGluThrLeuGluLeuSerGlySerArgLeu 335
DB 901 GTGATGTAAGTGTGAGGCTGCTTCTCTCAAGAAACCTGAGCTGTCTGGAGAAGATTTG 960
QY 336 GlnGln 337
DB 961 GAGCAG 966

```

RESULT 14

```

AF474987 969 bp DNA linear PRI 16-JUN-2003
LOCUS Homo sapiens G protein-coupled receptor SNR1 gene, complete cds.
DEFINITION AF474987
ACCESSION AF474987
VERSION AF474987.2 GI:31747894
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)

```

```

REFERENCE
AUTHORS Lembo, P.M., Grazzini, E., Groblewski, T., O'Donnell, D., Roy, M.-O.,
Zhang, D., Hoffert, C., Cao, J., Schmidt, R., Pelletier, M., Labarre, M.,
Gosselin, M., Fortin, Y., Banville, D., Shen, S., Strom, P., Payza, K.,
Dray, A., Walker, P. and Ahmad, S.
Proenkephalin A gene products activate a new family of sensory
neuron-specific GPCRs
JOURNAL Nat. Neurosci. 5 (3), 201-209 (2002)
MEDLINE 21853733
PUBMED 11850634

```

TITLE

```

JOURNAL 2 (bases 1 to 969)
REFERENCE Ahmad, S., Banville, D., Fortin, Y., Lembo, P.M. and O'Donnell, D.
AUTHORS Direct Submission
TITLE Submitted (25-JAN-2002) Molecular Biology, Astrazeneca, 7171
JOURNAL Frederick-Banting, Montreal, Quebec H4S 1Z9, Canada
COMMENT On Jun 16, 2003 this sequence version replaced gi:119338907.
FEATURES
source
1..969
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..969
/product="G protein-coupled receptor SNR1"
1..969
/note="GPCR"
/codon_start=1
/product="G protein-coupled receptor SNR1"

```

mRNA

```

CDS
1..969
/product="G protein-coupled receptor SNR1"
/note="GPCR"
/codon_start=1
/product="G protein-coupled receptor SNR1"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 4.2e-130      Length: 969
Score: 1675.00      Matches: 319
Percent Similarity: 99.69%      Conservative: 2
Best Local Similarity: 99.07%      Mismatches: 1
Query Match: 95.01%      Indels: 0
DB: 9      Gaps: 0

```

US-09-867-570-2 (1-337) x AF474987 (1-969)

QY 16 MetAspSerThrIleProValIleuGlyThrGluLeuThrProIleAsnGlyArgGlu 35
 DB 1 ATGGAATCCAACTCCAGCTTGGGTACAGAACTGACCAATCAACGAGCGTAGAGAG 60
 QY 36 ThrProCysTyrIleGlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
 DB 61 ACTCCTTGGCAGACGAGACCTGAGCTTACGGGGGTGACGTGCATCGTTCCCTTGTG 120
 QY 56 AlaLeuThrGlyAsnAlaValIleuThrPleuGlyCysArgMetArgArgAsnAla 75
 DB 121 GCGCTGACAGAAACGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuSerGlyValIle 95
 DB 181 GTCTCCATCACTCACTCCAACTGCTGCGCGCGCACTTCTTCTTACGGCGCACTT 240
 QY 96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
 DB 241 ATATGTTCCGCGTTAGCCCTCACTCAATATCCGCAATCCATCCCAAAATCTCAGTCT 300
 QY 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
 DB 301 GTGATGACCTTTCCTTCACTTATAGCCCTAAGCATGCTGAGCGCATCAGACGAGCGC 360
 QY 136 CysLeuSerIleLeuThrProIleTyrPheCysArgArgProArgTyrLeuSerSer 155
 DB 361 TGCCTGCTCCATCTGCGCCCATCTGCTACACGCGCGCCGCAATCTCTCATCTG 420
 QY 156 ValMetCysValIleuLeuThrAlaLeuSerLeuLeuArgSerIleLeuGluTyrPhePhe 175
 DB 421 GTATATGTATCTCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThr 195
 DB 481 TGTGACTCTCTGTTTACTGAGTGAATTTCTGTTGGTGAAGCATTCAGATTTCATTACA 540
 QY 196 IleAlaTyrPleuValPheLeuCysValIleuCysGlySerSerLeuValLeuVal 215
 DB 541 ATGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
 QY 216 ArgIleLeuCysGlySerArgGlyMetProLeuThrArgLeuTyrValIleLeuLeu 235
 DB 601 AGAATCTCTGTGATCCCGAAGATCCCGCTGACCGAGCTGACGATCTCTCTC 660
 QY 236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTyrAlaLeuPheSer 255
 DB 661 ACAGTGTGGTCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 256 ArgIleLeuAspTyrPheValLeuPheCysHisValIleValSerIlePheLeu 275
 DB 721 AGGATCCACTGAGTTGGAAAGTCTTATTTGTCATGTCATGATCTGATCTGATCTGCTG 780
 QY 276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGln 295
 DB 781 TCGGCTCTTAAACGACATGCAACCCCATCACTTCTTCTGCGGCTCTCTTAAAGCAG 840
 QY 296 ArgGlnAsnArgGlnAsnLeuValLeuValLeuGlnAlaGlnAlaGlnAlaPheProGlu 315
 DB 841 CGTCAAAATAGCGAGAACCTGAGCTGTCTCCAGAGGCTCTGACGAGACACCGCTGAG 900
 QY 316 ValAspGluGlyGlyTyrPleuProGlnGluThrLeuGluLeuSerGlySerArgLeu 335
 DB 901 GTGATGTAAGGTGGGGGTGCTTCTCAGAGAAACCTGAGAGCTGTGGGAGAGCAATTG 960
 QY 336 GluGln 337
 DB 961 GAGCAG 966
 RESULT 15
 LOCUS AR477199 969 bp DNA linear PAT 14-MAY-2004
 DEFINITION Sequence 4 from patent US 6696257.
 ACCESSION AR477199

VERSION AR477199.1 GI:47234473
 KEYWORDS Unknown.
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 969)
 AUTHORS Ahmed, S., Banville, D., Fortin, Y., Lembo, P., O'Donnell, D. and Shen, S.-H.
 TITLE G protein-coupled receptors from the rat and human
 JOURNAL Patent: US 6696257-A 4 24-FEB-2004.
 FEATURES Location/Qualifiers
 source 1..969
 /organism="unknown"
 /mol_type="genomic DNA"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,32e-127 Length: 969
 Score: 1642.00 Matches: 313
 Percent Similarity: 98.76% Conservative: 5
 Best Local Similarity: 97.20% Mismatches: 4
 Query Match: 93.14% Indels: 0
 DB: 6 Gaps: 0
 US-09-867-570-2 (1-337) x AR477199 (1-969)
 QY 16 MetAspSerThrIleProValIleuGlyThrGluLeuThrProIleAsnGlyArgGlu 35
 DB 1 ATGGAATCCAACTCCAGCTTGGGTACAGAACTGACCAATCAACGAGCGTAGAGAG 60
 QY 36 ThrProCysTyrIleGlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
 DB 61 ACTCCTTGGCAGACGAGACCTGAGCTTACGGGGGTGACGTGCATCGTTCCCTTGTG 120
 QY 56 AlaLeuThrGlyAsnAlaValIleuThrPleuGlyCysArgMetArgArgAsnAla 75
 DB 121 GCGCTGACAGAAACGGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuSerGlyValIle 95
 DB 181 GTCTCCATCACTCACTCCAACTGCTGCGCGCACTTCTTCTTACGGCGCACTT 240
 QY 96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
 DB 241 ATATGTTCCGCGTTAGCCCTCACTCAATATCCGCAATCCATCCCAAAATCTCAGTCT 300
 QY 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
 DB 301 GTGATGACCTTTCCTTCACTTATAGCCCTAAGCATGCTGAGCGCATCAGACGAGCGC 360
 QY 136 CysLeuSerIleLeuThrProIleTyrPheCysArgArgProArgTyrLeuSerSer 155
 DB 361 TGCCTGCTCCATCTGCGCCCATCTGCTACACGCGCGCCGCAATCTCTCATCTG 420
 QY 156 ValMetCysValIleuLeuThrAlaLeuSerLeuLeuArgSerIleLeuGluTyrPhePhe 175
 DB 421 GTATATGTATCTCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThr 195
 DB 481 TGTGACTCTCTGTTTACTGAGTGAATTTCTGTTGGTGAAGCATTCAGATTTCATTACA 540
 QY 196 IleAlaTyrPleuValPheLeuCysValIleuCysGlySerSerLeuValLeuVal 215
 DB 541 ATGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 600
 QY 216 ArgIleLeuCysGlySerArgGlyMetProLeuThrArgLeuTyrValIleLeuLeu 235
 DB 601 AGAATCTCTGTGATCCCGAAGATCCCGCTGACCGAGCTGTAGACGATCTCTCTC 660
 QY 236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTyrAlaLeuPheSer 255
 DB 661 ACAGTGTGGTCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720

```

QY 256 ArgIleHisLeuAspTrpIysValLeuPheCysHisValHisLeuValSerIlePheLeu 275
Db 721 AGGATCCACCTGGATTGGAAAGCTTATTGTGCAIGTGCACTAGTTCCATTTCCTG 780
QY 276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgIn 295
Db 781 TCCGCTCTTAACAGCAGTGCCCAACCCCATCATTTACTTCTTCGTGGGCTCCTTTAGGCAG 840
QY 296 ArgGlnAsnArgGlnAsnLeuLysLeuValLeuGlnArgAlaLeuGlnAspThrProGln 315
Db 841 CGTCAAAATAGGCAGAAACCTGAAAGCTGGTTCTCCAAAGGGCTCTGCAGACACGCTGAG 900
QY 316 ValAspGluGlyGlyTyrLeuProGlnGluThrLeuGluLeuSerGlySerArgLeu 335
Db 901 GTGGATGAAGGTGAGGGGTGGCTTCTCTCAGGAAACCTGAGACTGTGGGAGCAAAATTG 960
QY 336 GlnGln 337
Db 961 GAGCAG 966

```

Search completed: December 2, 2004, 00:13:55
 Job time : 4083 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 1, 2004, 21:28:03 ; Search time 486 Seconds
(without alignments)
3640.031 Million cell updates/sec

Title: US-09-867-570-2

Perfect score: 1763
1 MESKSKWVIRLGLFLSMDSTI.....EGGGWLPQETLBSGRLEQ 337

Scoring table:

BLASUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgn1.1/USPTO.spool.p/US09867570/runat.01122004.161122.6327/app.query.fasta_1.519
-DB=N_Geneseq.23Sep04 -QFMT=fasta -SUFFIX=p2n.rng -MINMATCH=0.1 -ICOPT=0
-LOOPEXT=0 -INIT=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdt
-LIST=45 -DOCLIGN=200 -THR SCORE=PCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=PCT -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US09867570 @CGN 1.1 708 @runat.01122004.161122.6327 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq.23Sep04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1763	100.0	2618	6	ABK52822 cDNA enco
2	1727.5	98.0	8622	6	ABK52823 Genomic D
3	1722.5	97.7	1369	10	ADCB6820 Human GPC
4	1721	97.6	1400	6	AAD33751 Human Mrp
5	1721	97.6	1400	8	ABZ42595 Human G p
6	1721	97.6	1400	12	ADH08534 DNA seque

7	1688	95.7	966	12	AD044603	Ad044603 Human HIT
8	1688	95.7	969	3	AA29811	AA29811 Human G p
9	1688	95.7	969	6	ABT04875	ABT04875 Human G p
10	1688	95.7	969	12	AD030080	Ad030080 Human GPC
11	1683	95.7	1683	10	ADF70583	Adf70583 Orphan re
12	1683	95.5	966	12	AD044605	Ad044605 Human HIT
13	1683	95.5	969	3	AA29812	AA29812 Human G p
14	1683	95.5	969	12	ADF29104	Adf29104 Human GPC
15	1663	94.3	969	3	AAA70342	AAa70342 Human nov
16	1642	93.1	969	2	AA210067	AAa210067 Human dor
17	1598	90.6	969	2	AA210068	AAa210068 Human dor
18	1557	88.3	909	8	ACC44771	ACC44771 Human G p
19	1416	80.3	1369	10	ADC86444	ADC86444 Human GPC
20	1416	80.3	2040	6	AAD33744	Ad33744 Human Mrp
21	1416	80.3	2040	12	ADH08519	Adh08519 DNA seque
22	1416	80.3	2040	12	AD036620	Ad036620 Human mas
23	1405	79.7	969	2	AA210069	AAa210069 Human dor
24	1399	79.4	912	10	ADC12765	Adc12765 Human GPC
25	1397	79.2	997	6	AD116635	Ad116635 Human NOV
26	1397	79.2	997	12	ADN42289	Adn42289 Human CDN
27	1390	78.8	1369	10	ADC86588	Adc86588 Human GPC
28	1388	78.7	1604	6	ABZ42602	Abz42602 Human Mrp
29	1388	78.7	1604	12	ADH08536	Adh08536 DNA seque
30	1387	78.7	1604	3	AA122008	AAa12208 Human G p
31	1387	78.7	969	10	ADC17729	Adc17729 Human TGR
32	1387	78.7	969	5	AAAD20943	AAa20943 Human G p
33	1383	78.4	969	5	AA079944	AA079944 Human CDN
34	1383	78.4	969	6	AAD27494	Ad27494 Human G-P
35	1383	78.4	969	6	ABZ42891	Abz42891 Human GPC
36	1383	78.4	969	8	ACC44772	ACC44772 Human G p
37	1383	78.4	969	8	ABZ81334	Abz81334 Human Dor
38	1383	78.4	969	10	AD61647	Ad61647 Human GPC
39	1383	78.4	969	11	ADL96465	Adl96465 Human G p
40	1383	78.4	969	12	AD030078	Ad030078 Human GPC
41	1383	78.4	1116	4	AAH49507	AAh49507 Human GTP
42	1383	78.4	969	2	AAZ10071	AAz10071 Human dor
43	1382	78.4	969	2	AAZ10070	AAz10070 Human dor
44	1382	78.4	969	3	AAA12204	AAa12204 Human G p
45	1367	77.5	969	3	AAA12204	AAa12204 Human G p

ALIGNMENTS

RESULT 1	ABK52822	
ID	ABK52822 standard; cDNA; 2618 BP.	
XX		
AC	ABK52822;	
XX		
DT	27-AUG-2002 (first entry)	
XX		
DE	cDNA encoding human G-protein coupled receptor (GPCR).	
XX		
KW	Human, G-protein coupled; receptor; GPCR; human protease;	
KW	human therapeutic protein; query sequence; search; gene; ss;	
KW	sequence database; non-human transgenic animal; gene therapy;	
KW	chromosome 3.	
XX		
OS	Homo sapiens.	
XX		
FT		
FT	Key	Location/Qualifiers
FT	CDS	447..1460
FT		/*tag= a
FT		/product= "Human G-protein coupled receptor (GPCR)"
XX		
FN	W0200234914-A1.	
XX		
PD	02-MAY-2002.	
XX		
PF	10-OCT-2001; 2001WO-US031592.	
XX		
PR	25-OCT-2000; 2000US-00695045.	
PR	31-MAY-2001; 2001US-00867570.	

XX (PEKE) PE CORP NY.
 PA
 XX
 PI Wei M, Zhao Q, Woodage T, Di Francesco V, Beasley EM;
 XX WPI: 2002-463360/49.
 DR P-PSDB: AAU97598.
 XX
 PT Novel isolated G-protein coupled receptor peptide useful for treating
 PT disorder characterized by absence of, in appropriate or unwanted
 PT expression of the receptor protein, and as immunogens to raise
 PT antibodies.
 PS
 PS Claim 4; Fig 1; 75pp; English.
 XX
 CC The present invention relates to a new G-protein coupled receptor (GPCR)
 CC peptide. The invention is useful for identifying a modulator of GPCR and
 CC for treating a disease or condition mediated by a human protease. The
 CC invention is also useful as models for the development of human
 CC therapeutics, for identifying therapeutic proteins, as targets for
 CC development of human therapeutic agents, and as query sequence to perform
 CC a search against sequence databases to, for e.g., identify other family
 CC members of related sequences. The vector of the invention is useful for
 CC producing a GPCR protein or peptide, for conducting cell-based assays
 CC involving the GPCR protein or its fragment, for identifying GPCR protein
 CC mutants whose functions are affected, and to produce non-human transgenic
 CC animals. The present nucleic acid sequence represents the human G-protein
 CC coupled receptor (GPCR) gene located on chromosome 3. This sequence
 CC encodes the human G-protein coupled receptor (GPCR) protein of the
 CC invention
 XX
 SQ Sequence 2618 BP; 655 A; 637 C; 546 G; 780 T; 0 U; 0 Other:
 Alignment Scores:
 Pred. No.: 2,876-155 Length: 2618
 Score: 1763.00 Matches: 337
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-867-570-2 (1-337) x ABK52822 (1-2618)
 QY 1 MetGluSerIySerSerTrpValIleArgLeuGlyPheLeuSerMetAspSerThrIle 20
 Db 447 ATGGAAGCAATCTCTCATGGCTCATCACTGGGTTTCTGAGCATGATTCACACATC 506
 QY 21 ProValIleuGlyThrGluLeuThrProIleAsnGlyArgGluGluThrProCysTyrIys 40
 Db 507 CCAGCTTGGGTACAGAACTGACACCAATCAAGGACGTAGAGAGACTCTTCTCAAG 566
 QY 41 GlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsn 60
 Db 567 CAGACCTGAGGCTTACAGGGGCTGAGCTGATCGTTCCCTGTCGCTGACAGGAAC 626
 QY 61 AlaValIleuThrPheLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIle 80
 Db 627 GCGGTGTGCTGTGCTCTCTGGCTGCCGATCGCAGAGAACCTCTCTCATCTACATC 686
 QY 81 LeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeu 100
 Db 687 CTCAACTGTGTGCGGCGCACTTCTCTCTTACGAGGCACTATATATGTCGCGTTA 746
 QY 101 ArgLeuIleAsnIleArgHisProIleSerIleSerIleLeuSerProValMetThrPhePro 120
 Db 747 CGCCCATCAATATATCGCATCCCATCTCAAAATCTCATGTCCTGTGATGACCTTCC 806
 QY 121 TyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeu 140
 Db 807 TACTTTATAGGCTTAAAGCATGCTGAGCGCATCAGACGAGGCTGCTGCTCATCTCG 866
 QY 141 TrpProIleTrpTyrHisCysArgArgProArgTyrLeuSerSerValMetCysValLeu 160

Db 867 TGGCCATCTGGNACACTGCGCGCCGACAGATACCTGTATCATGGTCATGTCCTG 926
 QY 161 LeuThrPalaleuSerLeuLeuArgSerIleLeuGluThrMetPheCysAspPheLeuPhe 180
 Db 927 CTCTGGGCGCTTCCCTGCTGGAGATTCCTGAGAGATGATGTCGACTTCTCTGTT 986
 QY 181 SerGlyAlaAspSerValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuVal 200
 Db 987 AGTGTGCTGATTCCTGTTGGTGTAAAGTCAGATTCATTACATTCAGCTGCTGTT 1046
 QY 201 PheLeuCysValIleuCysGlySerSerLeuValIleuLeuValArgIleuCysGly 220
 Db 1047 TTTTATGTGTGTTCTCTGTGGTCCAGCGCTGCTGTGTGTCAGGATTCCTGTGA 1106
 QY 221 SerArgIyMetProLeuThrArgLeuTyrValThrIleLeuThrValLeuValPhe 240
 Db 1107 TCCCGAAGATCCCGCTGACAGGCTGACGATCACTTCCTCAGTGTGCTTTC 1166
 QY 241 LeuLeuCysGlyLeuProPheGlyIleGlnThrPalaleuPheSerArgIleIleLeuAsp 260
 Db 1167 CTCTCTGTGGCTGCTGCTTGTGGCATTCAGTGGGCTCTTTCCAGATCCACTGAT 1226
 QY 261 TrpIyValIleuPheCysHisValHisIleuValSerIlePheLeuSerAlaLeuAsnSer 280
 Db 1227 TGGAAAGCTTATTTGTGATGTCATGATCATGTTTCATTTCCTGCTCGCTTAAACAGC 1286
 QY 281 SerIleAsnProIleIleTyrPhePheValGlySerPheAsnArgIleAsnArgGln 300
 Db 1287 AGTGCACACCCATCATTTACTTCTGTTGAGGCTCTTTAGGACAGCGTCAAAATAGGAG 1346
 QY 301 AsnLeuIyLeuValIleuGlnArgAlaLeuGlnAspThrProGluValAspGluGly 320
 Db 1347 AACCTGAAGCTGTTCTCTCAGAGGCTCTGACAGACACGCTGAGTGGATGAGAGTGA 1406
 QY 321 GlyTrpLeuProGlnGluThrLeuGluLeuSerGlySerArgLeuGluGln 337
 Db 1407 GGTGTGCTTCTCAGGAAACCTGTGAGCTGTGCGGACAGATTCAGACAG 1457
 RESULT 2
 ID ABK52823 standard; DNA; 8622 BP.
 AC ABK52823;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Genomic DNA encoding human G-protein coupled receptor (GPCR).
 KW Human; G-protein coupled; receptor; GPCR; human protease;
 KW human therapeutic protein; query sequence; search; gene; ds;
 KW sequence database; non-human transgenic animal; gene therapy;
 KW chromosome 3.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS /tag= a 4300..8496
 FT /product= "Human G-protein coupled receptor (GPCR)"
 FT exon /tag= b 4300..4319
 FT /number= 1
 FT intron /tag= c 4320..7502
 FT /number= 1
 FT exon /tag= 1 7503..8496
 FT /tag= d
 FT /number= 2
 XX
 PD 02-MAY-2002.
 XX
 XX WO200234914-A1.

```

PF 10-OCT-2001; 2001WO-US031592.
XX
XX 25-OCT-2000; 2000US-00695045.
PR 31-MAY-2001; 2001US-00867570.
XX
XX (PEKE ) PR CORP NY.
XX
PI Wei M, Zhao Q, Woodage T, Di Francesco V, Beasley EM;
XX WPI; 2002-463360/49.
DR P-PSDB; AAU97598.
XX
XX Novel isolated G-protein coupled receptor peptide useful for treating
PT disorder characterized by absence of, in appropriate or unwanted
PT expression of the receptor protein, and as immunogens to raise
PT antibodies.
XX
PS Claim 4; Fig 3; 75pp; English.
XX
XX The present invention relates to a new G-protein coupled receptor (GPCR)
CC peptide. The invention is useful for identifying a modulator of GPCR and
CC for treating a disease or condition mediated by a human protease. The
CC invention is also useful as models for the development of human
CC therapeutics, for identifying therapeutic proteins, as targets for
CC development of human therapeutic agents, and as query sequence to perform
CC a search against sequence databases to, for e.g., identify other family
CC members of related sequences. The vector of the invention is useful for
CC producing a GPCR protein or peptide, for conducting cell-based assays
CC involving the GPCR protein or its fragment, for identifying GPCR protein
CC mutants whose functions are affected, and to produce non-human transgenic
CC animals. The present nucleic acid sequence represents the human G-protein
CC coupled receptor (GPCR) gene located on chromosome 3. This sequence
CC encodes the human G-protein coupled receptor (GPCR) protein of the
CC invention
XX
XX Sequence 8622 BP; 2684 A; 1921 C; 1679 G; 2338 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 3,04e-151 Length: 8622
Score: 1727.50 Matches: 332
Percent Similarity: 96.81% Conservative: 2
Best Local Similarity: 96.23% Mismatches: 2
Query Match: 97.99% Indels: 9
DB: 6 Gaps: 1
US-09-867-570-2 (1-337) x ABK52823 (1-8622)
QY 2 GluSerLysSerSerTrp-----ValIleArgLeuGly 12
DB 7459 CAGAGATCAAAACAGCTGGATACATCTGGTTCTGTTCCAGAGGTCATCAGCTGG 7518
QY 13 PheLeuSerMetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGly 32
DB 7519 TTTCTGAGCATGGATTCAACCATCCAGTCTGGGTACAGAACATCAACCATCAAGGA 7518
QY 33 ArgGluGluThrProCysTyrTyrGlnThrLeuSerPheThrGlyLeuThrCysIleVal 52
DB 7579 CGTGAGAGAGACTCTTCTCTCAAGACAGACCTTGAGTTCAAGGGGCTGAGCTGATGCTT 7638
QY 53 SerLeuValAlaLeuThrGlyAspAlaValValLeuThrLeuLeuGlyCysArgMetArg 72
DB 7639 TCCCTTGTCCCGCGAGAGAAAGCGGTTGTCTGTGGCTCCGCGGCTCCGATGCGC 7698
QY 73 ArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaIleAspPheLeuPheLeuSer 92
DB 7699 AGGAGCGCTCTCTCACTCATCTCACTGCTGCGGCGCATCTTCTTCTTACG 7758
QY 93 GlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIle 112
DB 7759 GGGCCACATTAATGTTGCGGTACGCTCATCAATATCCGCCATCCATCTCCAAATC 7818
QY 113 LeuSerProValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSer 132

```

```

DB 7819 CTGAGTCTGTGATGAGACTTTCCCTACTTATAGGCTTAAGCATGCTGAGCGCATACG 7878
QY 133 ThrGluArgCysLeuSerIleLeuThrProIleTyrHisCysArgArgProArgTyr 152
DB 7879 ACCGAGGCTGCGCTGTCATCTGTGGCCCATCTGGTACACCATGCGCCGCCAGATAC 7938
QY 153 LeuSerSerValMetCysValIleLeuThrPAlaIleSerLeuLeuArgSerIleGlu 172
DB 7939 CTGTCAATCGGTCAATGTGTGTCTGCTGCGGCTCTGTCTCCGTGCGGATATCTGGAG 7998
QY 173 TrpMetPheCysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAsp 192
DB 7999 TGGATGTTCTGTATCTTCTTCTTATAGGTCGTCATCTTCTTGTGTAACGTCATGAT 8058
QY 193 PheIleThrIleAlaTyrLeuValPheLeuCysValValLeuCysGlySerSerLeuVal 212
DB 8059 TTCATTPACATCCGTCGTCGCTGTTTATTTATGTGTGTTCTGTGGGTACACCTGCTC 8118
QY 213 LeuLeuValAlaArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValThr 232
DB 8119 CTGCTGTCAAGATTCCTGTGTGATCCCGAAGATGCGCTGACACAGCTGTACGTGACC 8178
QY 233 IleLeuLeuThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTyrAla 252
DB 8179 ATCTCTCTCAAGATGCTGCTGCTTCTTCTCTGTGCTGCTGCTTGGCATTCAGTGGGCC 8238
QY 253 LeuPheSerArgIleHisLeuAspTyrPLeuValLeuPheCysHisValHisLeuValSer 272
DB 8239 CTGTTTCCAGATTCACCTGATGGAAGTCTTATTTGTGTCATGTCATGATGATTTC 8298
QY 273 IlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySer 292
DB 8299 ATTTTCTGTCCTCTCTTAAACAGACAGTCCACATTCATTCCTTCTGCGGCTCC 8358
QY 293 PheArgGlnArgGlnAsnArgGlnAsnLeuValLeuGlnIleArgAlaLeuGlnAsp 312
DB 8359 TTTAGCAGCGCTCAAAATAGGCAAGAACCTGATGTTTCCAGAGGCTCTGACAGAC 8418
QY 313 ThrProGluValAspGluGlyGlyTyrPLeuProGlnGluThrLeuGluLeuSerGly 332
DB 8419 ACCGCTGAGGTGATGAAAGGTGAGGCTTCTTCCAGAAACCTCGAGCTGTGCGGA 8478
QY 333 SerArgLeuGluGln 337
DB 8479 AGCAGATTGGAGCAG 8493
RESULT 3
ADCB6820
ID ADCB6820 standard; DNA; 1369 BP.
XX
XX ADCB6820;
AC 01-JAN-2004 (first entry)
XX
XX Human GPCR gene SEQ ID NO:1273.
XX
XX de; gene; human; GPCR;
XX
XX guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX
XX Homo sapiens.
OS
XX
XX BP1270724-A2.
XX
XX 02-JAN-2003.
XX
XX 18-JUN-2002; 2002EB-00013517.
XX
XX 18-JUN-2001; 2001JP-00246789.
XX
XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
XX Suwa M, Asai K, Akiyama Y, Aduzatani H;
PI

```

XX MPI: 2003-315783/31.
DR P-PsDB; ADC686821.

PT New polynucleotide, useful for preparing a composition for treating a
PR patient in need of increased or suppressed activity or expression of the
PT guanosine triphosphate-binding protein coupled receptor.

XX
XX
PS Claim 1; SEQ ID NO 1273; 28pp; English.

CC The invention relates to a novel polynucleotide encoding a guanosine
CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
CC the invention may have a use in gene therapy. The polynucleotide and
CC polypeptide are useful for preparing a composition for treating a patient
CC in need of increased or suppressed activity or expression of the
CC guanosine triphosphate-binding protein coupled receptor. The
CC polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
CC invention.

XX
XX

SQ Sequence 1369 BP; 272 A; 372 C; 340 G; 385 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	7,64e-152	Length:	1369
Score:	1722.50	Matches:	331
Percent Similarity:	96.81%	Conservative:	3
Best Local Similarity:	95.94%	Mismatches:	2
Query Match:	97.70%	Indels:	9
DB:	10	Gaps:	1

US-09-867-570-2 (1-337) x ADC686820 (1-1369)

OY 2 GluserylserSerP-----ValleargLeugly 12
Db 132 CAGAGATCAAAACGGCTGGTGCATCATCGTTGTCCTTCCAGGGTCACTCAGACTGAGCTGGG 191
OY 13 PheleuseMerapSerThrileProvalLeuglyThrguleuthrProilleasngly 32
Db 192 TTTCATGACATGATTCACCACTCCAGCTTGGGTACAGAACGTACAACCATACGA 251
OY 33 ArgJugluThrProCysTyrylvgInthrleuserPhethrgrjyleuthrCysileval 52
Db 252 CGGAGAGAGACTCTTGCTACAGAAGACCCTGAGCTTCAAGGGGCTGACGTGATCGTT 311
OY 53 SerleuvalAlaleuthrgrjyaenAlavalLeuthrpleuleuglyCybaRmeArg 72
Db 312 TCCCTTGGCGCGGTGACAGAAACGGGTGTGCTCTGCTCGCTCGGGCTGCCGATCGC 371
OY 73 ArgAsnAlaValSerIetyrIleleuAsnleuValAlalaAppPheleuPheleuSer 92
Db 372 AGGAGCGCTGTCTCCATCTCACCTTAAGCTGTGGGGCCGACTTCTCTTCTTAGC 431
OY 93 GlyHisIlelleCysSerProleuArgleulleanllearghisProilleserVale 112
Db 432 GGCCACATTATATGTTGCGCGTTAAGCCTCATCAATACGCCATCCCATCCATAATC 491
OY 113 LeuSerProvalMetThrPheProTyrrPheIlleglyeuserMetleuserAlleser 133
Db 492 CTCAGTCCGTGTGATGACCTTCTCCACTTATAGCCCTAAGCATGCTGAGGCCATCAGC 551
OY 133 ThrGluArCybaSerIleleuthrProilleftPrYrhIsyBAARGARProArgTYr 151
Db 552 ACCGACCGCTGCTTCATCTCTGTGCCCCCATCTGTATACCATCCGCCGCCCAAGTAC 611
OY 153 LeuSerSerValMetCysValleuLeuThrAlaleuSerleuAurgserrileleuglu 171
Db 612 CTGTATCATAGTATGATGTGTCGTCTGGGCCCTGTCCCTGCGGGAAGATCTGGAG 671
OY 173 TrpMetPheCysaspPheleupheserGlyAlalaspSerValITPyrgIuthSerASP 191
Db 672 TGATATTTCTGTACTTCTCTGTAGAGGTCTAATCTGTTTGSTGTGAACCTCAGAT 731
OY 193 PheIIleThrIleAlatPheleuValPheleuCysValValleucysglYserSerleuVal 211

Db	732	TTCAATTAAACAATCGCGTGGCTGTTTATGTGAGTTCCTGTGGGTCACAGCTGATC	791
Qy	213	LeuLeuValaArgjleLeuCyeglySerArgjySkeTProLeuThrArgLeuTyValThr	232
Db	792	CTGCTGGACAGAGATTCTCTGTGGATCCCGAAGATGCCGCTGACCAAGGCTGTAGCTGACC	851
Qy	233	lleLeuLeuThrValleuValaPheLeuLeuCyeglyLeuProPheglylleGlnTrpala	252
Db	852	ATCCTCCCTCACAGTCTGTGCTTCTCTCTCTGTGGCCCTGCCCTTTGGCATTCAGTGGGCC	911
Qy	253	LeuPheSerArgjlelleLeuAspTrpIysValleuPheCySHIsvaIleLeuValSer	272
Db	912	CTGTTTTCAAGATCACCTGGATTTGAAGATCTTATTTGTCAATGTGCATCTAGTTTCC	971
Qy	273	llePheLeuSerArgjleLeuAsnSerSerArgjleLeuProIleIleTyTrpPheValGlySer	292
Db	972	ATTTCCTGTGTCCGCTCTTACACAGCGTGCACAACTTCATCTTCTTGTGGGCTCC	1031
Qy	293	PheArgjleuArgjleuAsnArgjleuAsnLeuIysleuValleuGlnArgjleuGlnAsp	312
Db	1032	TTTAGGCGAGCGTCAAAATAGGCAGAACCTGAAAGCTGTCTTCCAGAGGGCTCTGACGAC	1091
Qy	313	ThrProGlnValaAspGlnGlyGlyGlyTrpLeuProGlnGlnIuThrLeuGlnLeuSerGly	332
Db	1092	ACGCTGAGGTGATGAGAGTGTGAGGGGTGGCTTCTCCTCAGAAACCTGTGACCTGTGGGA	1151
Qy	333	SerArgLeuGlnGln 337	
Db	1152	AGCAGATTGGAGCAG 1166	
RESULT 4			
ADD33751	ID	ADD33751 standard; DNA; 1400 BP.	
AC	ADD33751;		
DT	01-JUL-2002	(first entry)	
DE	Human Mrx3 (mas-related gene) DNA.		
KW	Human; mas-related gene; G-protein coupled receptor; drg-12 protein;		
OS	Homo sapiens.		
Key	Location/Qualifiers		
FT	332..1300		
FT	/*tag= a		
FT	/product= "Human Mrx3 protein"		
PN	WO200183555-A2.		
PD	08-NOV-2001.		
PF	04-MAY-2001; 2001MO-US014519.		
PR	04-MAY-2000; 2000US-0202027P.		
PR	01-AUG-2000; 2000US-0222344P.		
PR	03-NOV-2000; 2000US-00704707.		
PR	19-APR-2001; 2001US-0285493P.		
PA	(CALY) CALIFORNIA INST OF TECHNOLOGY.		
PI	Anderson DJ, Dong X, Zylka M, Han S, Simon M;		
DR	WPI: 2002-171346/22.		
DR	P-PSDB; AA021296.		
XX	Isolated polypeptide, Mrx, which is a G-protein coupled receptor and an		
PT	isolated polypeptide, drg-12, which is also a receptor, useful for		
PT	identifying agonists or antagonists for treating pain.		
PS	Disclosure; Page 128-129; 185pp; English.		

XX The invention relates to Mrx (mas-related gene) protein, which is a G-protein coupled receptor and drg-12 protein, which is a receptor. The invention is useful for identifying compounds that bind to it, especially agonists or antagonists. Administration of an agent (e.g. the identified agent) that increases the expression of Mrx in a mammal may be used for treating impaired sensory perception in a mammal, especially pain. The agent may also be useful for treating impaired sensory perception in a mammal. The present sequence is human Mrx3 DNA

Sequence 1400 BP; 276 A; 375 C; 363 G; 386 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,09e-151	Length:	1400
Score:	1721.00	Matches:	330
Percent Similarity:	99.40%	Conservative:	0
Best Local Similarity:	99.40%	Mismatches:	2
Query Match:	97.62%	Indels:	0
DB:	6	Gaps:	0

US-09-867-570-2 (1-337) x MAB33751 (1-1400)

Qy 6 SerTrpValIleAryLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThr 25
 Db 302 TCCAGGGTCAACGACTGGGGTTCTGACATGATTCACACATCCAGCTTGGGTAACA 361
 Qy 26 GluLeuThrProIleAengIyArgIuGluThrProCysTyrIlyGlnThrLeuSerPhe 45
 Db 362 GAACGTACACCAATCAACGACGAGAGACTCTTGCTCAAGACAGACCTCAGCTTC 421
 Qy 46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAmaAlaValLeuTrp 65
 Db 422 ACCGGGCTGACGAGCATCGTTCCCTTGTGCGGTGACAGAAACGGGGTGTCTCTGG 481
 Qy 66 LeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIleAsnLeuValAla 85
 Db 482 CTCTGGGGCTGCCGATGCGGAGAACGCTGTCTCCATCTACATCTCCACCTGTCGG 541
 Qy 86 AlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleAsnIle 105
 Db 542 GCGGACTTCTCTCTCTTAGCGGCCACATTAATGTTCCGCGGTACGCTCATCATATC 601
 Qy 106 ArgHisProIleSerIyIleLeuSerProValMetThrPheProIyIleGlyLeu 125
 Db 602 CGCCATCCCATCTCCAAATCTCTGATCTGTGATGACCTTCCCTACTTATAGGCTTA 661
 Qy 126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTrpProIleTrpTyr 145
 Db 662 AGCATGCTGAGCCATCAAGACCGAGCGCTGCTGTCCATCTGTGAGCCCATCTGATAC 721
 Qy 146 HisCysArgArgProArgTyrLeuSerSerValMetCysValLeuLeuTrpAlaLeuSer 165
 Db 722 CACTGCGCGGCCCCAGATACCTGTCAACGCTGATGATGTCCTGCTGCGCCCTGCTGC 781
 Qy 166 LeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeuPheSerGlyAlaAspSer 185
 Db 782 CTGTGTGGAGATACCTGAGATGATGTTCTGTACTTCCCTGTTAGTGCGCTGATTTCT 841
 Qy 186 ValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuValPheLeuCysValVal 205
 Db 842 GTTTGGTGTAAACGTACGATTTTCAATACCGCGTGGCTTTTATATGATGATG 901
 Qy 206 LeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgIyMetPro 225
 Db 902 CTCTGTGGGTCAAGCTGTCTCTGTGATGATTTCTGTGATCCCGGAAGATGCGG 961
 Qy 226 LeuThrArgLeuTrpValThrIleLeuLeuThrValLeuValPheLeuCysGlyLeu 245
 Db 962 CTGACCGAGGCTGACGATCCTCTCAAGATGCTGCTTCTCTCTGAGGCTG 1021
 Qy 246 ProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAspTrpIyValLeuPhe 265
 Db 1022 CCTTTGGCATTCAGTGGGCCCTGTTTCCAGATCCACCTGATGGAAAGTCTTATTT 1081

Qy 266 CysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIle 285
 Db 1082 TGTCAATGTGATCTAGATTTCATTTCTCTGTCGCTCTTAAACAGCATGCGCAACCCATC 1141
 Qy 286 IleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgGlnAsnLeuValLeuVal 305
 Db 1142 ATTACTTCTTGTGGCTCTCTTTAGCAGCGTCACAAATAGCGAACCCTGAACTGGTT 1201
 Qy 306 LeuGlnArgAlaLeuGlnAsnArgTrpGluValAlaAspGluGlyGlyTyrIleProGln 325
 Db 1202 CTCAGAGGGCTCTGACAGACAGCCCTGAGGTGATTAAGTGGAGGTGGCTTCTCAG 1261
 Qy 326 GluThrLeuGluLeuSerGlySerArgLeuGluGln 337
 Db 1262 GAAACCTGAGAGCTGTGCGGAGACAGATGAGAGAG 1297
 RESULT 5
 AB242595
 ID AB242595 standard; DNA, 1400 BP.
 AC AB242595;
 XX
 DT 04-MAR-2003 (first entry)
 DE
 XX Human G protein-coupled receptor Mrx3 nucleotide SEQ ID NO:673.
 XX
 XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
 KW G protein-coupled receptor modulator; antibody; immune-related disease;
 KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KW immunological-related cell proliferative disease; autoimmune disease;
 KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KW osteoporosis; cardiovascular; inflammation; Crohn's disease; diabetes;
 KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KW ulcer; gene; ds.
 OS
 XX Homo sapiens.
 XX
 PN W0200261087-A2.
 XX
 PD 08-AUG-2002.
 XX
 XX 19-DEC-2001; 2001WO-US050107.
 PF
 XX 19-DEC-2000; 2000US-0257144P.
 PR
 XX (LIFE-) LIFESPAN BIOSCIENCES INC.
 PA
 XX Burner GC, Roush CL, Brown JP;
 PI
 XX WPI; 2003-046718/04.
 DR P-PSDB; ABP81750.
 DR
 XX New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR), useful for diagnosing and designing drugs for treating conditions
 PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
 PT autoimmune diseases.
 XX
 PS Disclosure; Fig 1; 523pp; English.
 XX
 XX The present invention describes antigenic peptides (I) comprising: (a)
 CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular G
 CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity or
 CC avidity for a particular GPCR. (1) can be used as GPCR modulators and in
 CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
 CC antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCR. The antigenic peptides for

CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related diseases, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 CC hypotension, renal disorders, rheumatoid arthritis, cirrhosis, ulcers, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
 CC exemplification of the present invention

XX Sequence 1400 BP; 276 A; 375 C; 363 G; 386 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.09e-151	Length:	1400
Score:	1721.00	Matches:	330
Percent Similarity:	99.40%	Conservative:	0
Best Local Similarity:	99.40%	Mismatches:	2
Query Match:	97.62%	Indels:	0
DB:	8	Gaps:	0

US-09-867-570-2 (1-337) x ABZ42595 (1-1400)

QY 6 SerTPValIleArgLeuGlyPheLeuSerMetLeuSerThrIleProValLeuGlyThr 25
 Db 302 TCCAGGGTCACACGAGCTGGGTTTCTGACATGATTCACACCTCCAGCTTGGGATAC 361
 QY 26 GluLeuThrProIleAsnGlyArgGluGluThrProCysTyrLeuGlnThrLeuSerPhe 45
 Db 362 GAATGACACCAATCAACGAGCGTAGAGACTCTTGCTACAGAGAACCTGAGACTTC 421
 QY 46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValValLeuTyr 65
 Db 422 ACGGGGCTGACGTGATCGTTTCCCTTGTCGGCTACAGGAACCGGTGTGCTCG 481
 QY 66 LeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAla 85
 Db 482 CTCTGGGGCTGCGCATGCGAGGAGAACGCTGCTCATCACTCAACCTGCTCGCG 541
 QY 86 AlaAspPheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleAsnIle 105
 Db 542 GCCGCTTCTCTTCTTACCGGCGCATTAATATGTTGCGGTTAGCCCTCATCAATATC 601
 QY 106 ArgHisProIleSerIleLeuSerProValMetThrPheProTyrPheIleGlyLeu 125
 Db 602 CGCCATCCCATCTCCAAATTCCTGATCGTATGATACCTTTCCCTTATATAGGCTTA 661
 QY 126 SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTyrProIleTyr 145
 Db 662 AGCAGCTGAGCGCATCAAGCACCGAGCGCTGCTTCCATCTTGCGCCCATCTGTTAC 721
 QY 146 HisCysArgArgProArgTyrLeuSerSerValMetCysValLeuLeuTyrPalaLeuSer 165
 Db 722 CACTCCCGCGCCGACGATACCTGTCATCGATGATGATGCTGCTCGGCGCTGCTTC 781
 QY 166 LeuLeuArgSerIleLeuGluTyrPheCysAspPheLeuPheLeuSerGlyAlaAspSer 185
 Db 782 CTGCTGCGGAGATCTCTGAGAGTGATGTTCTGTGACTTCTGTTAGTGTCTGATTTCT 841
 QY 186 ValTyrCysGluThrSerAspPheIleThrIleAlaTyrPheValPheLeuCysValVal 205
 Db 842 GTTGTGTGGAACGTCATATTCATTACATGCGTGGGTGTTTATATGTGTGT 901
 QY 206 LeuCysGlySerSerIleuValLeuValArgIleLeuCysGlySerArgIleMetPro 225
 Db 902 CTCTGTGGTCCAGCTGTGCTGCTGCTGAGGATTCCTGTGTGATCCCGGAAGAGCGC 961
 QY 226 LeuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeu 245

Db 962 CTGACAGAGCTGTACGTGACCATCTCTCCACAGAGCTGTGCTTCTCTCTGTGGCTTG 1021
 QY 246 PropheGlyIleGlnTyrPalaLeuPheSerArgIleHisIleAspTyrPalaLeuPhe 265
 Db 1022 CCGTTTGGATTCATGATGGGCGCTGTTCCTCAGAGATCCAGCTGATTTGAAAGTCTTATTT 1081
 QY 266 CysHisValHisIleuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIle 285
 Db 1082 TGTGATGTGCATCTAGTTTCCATTTTCTGTGCGCTTTACAGAGAGGCCAACCCATTC 1141
 QY 286 IleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgGlnAsnLeuIleVal 305
 Db 1142 ATTACTTCTTGTGGGCTCTTTAGGCAAGCGTCAAAATAGGCAAGACCTGAAGCTGGTT 1201
 QY 306 LeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGlyTyrLeuProGln 325
 Db 1202 CTCAGAGGGCTCTCAGAGACGCGCTAGGTGAAGGTGAGGTGGCTTCTGAG 1261
 QY 326 GluThrLeuGluLeuSerGlySerArgLeuGluGln 337
 Db 1262 GAACCTTGAGCTGTGGGAGACGATTTGAGCAG 1297
 RESULT 6
 ID ADH08534 standard; DNA; 1400 BP.
 AC ADH08534;
 DT 25-MAR-2004 (first entry)
 DE DNA sequence MrgX3.
 XX mas-related gene D; MrgD; Analgesic; Vulnerary; Ophthalmological;
 KM sensory perception; Glaucoma; Mrg; ds.
 OS Mus musculus.
 PN M02004003133-A1.
 XX 08-JUN-2004.
 PD 13-MAY-2003; 2003WO-US015004.
 PF 26-JUN-2002; 2002US-00183116.
 PR (CALY) CALIFORNIA INST OF TECHNOLOGY.
 PA Anderson DJ, Dong X, Zylka M, Han S, Simon MI;
 PI WPI; 2004-083025/08.
 DR P-PSDB; ADH08535.
 PT New mas-related gene D polypeptides, useful as therapeutics or in
 PT identifying agonists or antagonists that alter pain perception in a
 PT mammal for treating impaired sensory perception, e.g. chronic intractable
 PT pain or neuropathic pain.
 XX disclosure; SEQ ID NO 30; 220pp; English.
 XX The present invention relates to an isolated mas-related gene D (MrgD)
 CC polypeptide. The MrgD polypeptides are useful as therapeutics or for
 CC identifying compounds, i.e. agonists or antagonists, that alter pain
 CC perception in a mammal. The compounds are useful for treating impaired
 CC sensory perception, e.g. chronic intractable pain or neuropathic pain,
 CC promoting wound healing, restoring normal sensitivity following injury,
 CC or treating ocular conditions, particularly those associated with
 CC pressure such as glaucoma. The Mrg genes or proteins may be used as
 CC molecular probes for the detection of cells or tissues related to or
 CC involved with sensory perception. The present sequence represents a MrgA
 CC (Mrg subfamily) encoding sequence.
 XX Sequence 1400 BP; 276 A; 375 C; 363 G; 386 T; 0 U; 0 Other;


```

Db      1 ATGATTACACATCCAGTCTTGGGTACAGAACTACACCATCAACGAGCGTGAAGAG 60
Qy      36 ThrProCysTyrIleuGlnThrIleuSerPheThrGlyLeuThrCysIleValSerIleuVal 55
Db      61 ACTCCTTCTCTACACAGACCCCTGAGCTTCACGGGGCTGACGCGATCTCTTCCCTTGTGC 120
Qy      56 AlaIleuThrGlyAsnAlaValIleuThrPheLeuGlyCysArgMetIleArgAsnAla 75
Db      121 GCGCTGACAGGAAAGCGGTGTGTCTCTGCTCTGAGCTCTGGCTGCGGACGCGAGAAAGCT 180
Qy      76 ValSerIleTyrIleuAsnIleuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95
Db      181 GTCCTCATCTACATCTCAACCTGGTGGCGGCGGACTTCTCTCTTCTTACGCGGCGACATT 240
Qy      96 IleCysSerProLeuArgIleuIleAsnIleArgHisProIleSerIleLeuSerPro 115
Db      241 ATATGTTGCGCCCTTACGCTCATCATCATATCCGCGCATCTCCAAAATCTCAAGTCTCT 300
Qy      116 ValMetThrPheProTyrPheIleGlyLeuSerMetIleuSerAlaIleSerThrGlyArg 135
Db      301 GTGATGACCTTCTCTACTTATATAGCTTACAGATCTAGCGCCATACAGACCGAGCGC 360
Qy      136 CysIleuSerIleLeuThrProIleTyrHisCysArgArgProArgTyrIleuSerSer 155
Db      361 TGCGTGTCCATCTGTGGCCCATCTGTGACACTGCGCGCGCCGACATACCTGTATCG 420
Qy      156 ValMetCysValIleuLeuThrPalaIleuSerIleuArgSerIleuGluTyrMetPhe 175
Db      421 GTCAATGTGTCTGTCTGTGGCGCCCTGCTCCCTGCTCGAGTACCTGTGAGTGAATGTC 480
Qy      176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThr 195
Db      481 TGTGACTTCTCTTATGATGTGTCTGTATCTGTGTGTGAAAGTCATGATTTTCATTACA 540
Qy      196 IleAlaTyrPheValPheLeuCysValIleuCysGlySerSerIleuValIleuVal 215
Db      541 ATCGGTGGCTGGTCTTTTATGTGTGTGTCTGTGTGGTCCAGCGCTGTGTGTGTC 600
Qy      216 ArgIleuLeuCysGlySerArgIleuMetProLeuThrArgIleuTyrValThrIleuLeu 235
Db      601 AGGATTTCTTGTGATCCCGAGAGATGCGCTGACCGATGACGATCTCTCTC 660
Qy      236 ThrValIleuValPheLeuLeuCysGlyLeuProPheGlyIleGlnThrPalaIleuPheSer 255
Db      661 ACAGTGTGTGTCTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Qy      256 ArgIleuIleuAspTyrPheValIleuPheCysHisValHisIleuValSerIlePheLeu 275
Db      721 AGGATCCACCTGGATTGAAAGCTTATTTTGTCAATGTGATCTCAATTTTCTCTG 780
Qy      276 SerAlaIleuAsnSerSerIleuAsnProIleIleTyrPhePheValGlySerPheArgGln 295
Db      781 TCCGCTCTTAAACGACGTCACCAACCCCATCATTTACTCTTCTGTGGCTCTCTTTAGCAG 840
Qy      296 ArgGlnAsnArgGlnAsnIleuValIleuValIleuGlnArgAlaIleuGlnAsnThrProGlu 315
Db      841 CGTCAAAATAGGCAAGAACTGAAAGCTGTCTTCAGAGGCTGTGACGAGCAAGCGCTGAG 900
Qy      316 ValAspGluGlyGlyIleTyrPheLeuProGlnIleThrIleuGluLeuSerGlySerArgLeu 335
Db      901 GTGGATGAAAGTGTAGGTGTGTCTCTCTCAGAAACCTGTGAGCTGTGCGGAGACAGATTG 960
Qy      336 GluGln 337
Db      961 GAGCAG 966

```

RESULT 8
 ID AAA29811
 XX AAA29811 standard; cDNA; 969 BP.
 AC AAA29811;
 XX

```

DT      18-AUG-2000 (first entry)
XX
DE      Human G protein-coupled receptor hH17T213 encoding cDNA SEQ ID NO:3.
XX
KW      Human; G protein-coupled receptor; hippocampus; diagnosis; screening;
XX      genetic disease; cellular function regulation; ss.
XX
OS      Homo sapiens.
XX
FH      Key Location/Qualifiers
FT      CDS 1..969
FT      FT /*tag= a
FT      PN /product= "G protein-coupled receptor"
XX
XX      MO200020455-A1.
XX
XX      13-APR-2000.
XX
XX      30-SEP-1999; 99WO-JP005366.
XX
XX      01-OCT-1998; 98JP-00279535.
XX
XX      (TAKE ) TAKEDA CHEM IND LTD.
XX
XX      Watanabe T, Terao Y, Matsui H;
XX      MPI; 2000-303747/26.
XX      P-PSDB; AA90761.
XX
XX      Human-derived G protein-coupled protein and encoding nucleic acid, useful
XX      PT e.g. in determining ligands and treatment of diseases associated with
XX      PT dysfunction of the protein.
XX
XX      Claim 6; Page 93-94; 97pb; Japanese.
XX
XX      The present sequence encodes a human-derived G protein-coupled protein
XX      CC designated hH17T213, which is isolated from the human hippocampus. The G
XX      CC protein-coupled receptor can be used for preventing, treating and
XX      CC diagnosing genetic diseases associated with G protein-coupled protein,
XX      CC and for regulating cellular functions. The protein can be used to prevent
XX      CC and treat disorders associated with G protein-coupled protein gene
XX      CC dysfunction. It can also be used to identify G protein-coupled protein
XX      CC ligands and generating antibodies and antisera against the protein. It is
XX      CC also useful in constructing recombinant receptor protein expression
XX      CC systems, developing receptor-binding assay systems and screening drug
XX      CC candidates, and can be used as a probe in the genetic diagnosis of G
XX      CC protein-coupled protein disorders
XX
XX      Sequence 969 BP; 171 A; 280 C; 250 G; 268 T; 0 U; 0 Other;
XX
XX
XX      Alignment Scores:
XX      Pred. No.: 8,34e-149 Length: 969
XX      Score: 1688.00 Matches: 322
XX      Percent Similarity: 100.00% Conservative: 0
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 95.75% Indels: 0
XX      DB: 3 Gaps: 0
XX
XX      US-09-867-570-2 (1-337) x AAA29811 (1-969)
XX
Qy      16 MetAspSerThrIleProValIleuGlyThrGluLeuThrProIleAsnGlyArgGluGlu 35
Db      1 ATGATTACACATCCAGTCTTGGGTACAGAACTGACACCAATCAACGAGCGTGAAGAG 60
Qy      36 ThrProCysTyrIleuGlnThrIleuSerPheThrGlyLeuThrCysIleValSerIleuVal 55
Db      61 ACTCCTTCTCTACACAGACCCCTGAGCTTCACGGGGCTGACGCGATCTCTTCCCTTGTGC 120
Qy      56 AlaIleuThrGlyAsnAlaValIleuThrPheLeuGlyCysArgMetIleArgAsnAla 75
Db      121 GCGCTGACAGGAAAGCGGTGTGTCTCTGCTCTGAGCTCTGGCTGCGGACGCGAGAAAGCT 180
Qy      76 ValSerIleTyrIleuAsnIleuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95

```

```

Db      181 GTCTCCATCTACATCTCACTGCTGCGCCGCACTTCTCTTCTTACGCGCACTT 240
Qy      96  ILeCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
Db      241 ATATGTTGCGCGGTACCGCTCATCATATATCGGCATCTCCAAATCTCAGTCC 300
Qy      116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
Db      301 GTATGACCTTTCCCTTACTTATAGGCTTAAGCATGCTGACGCGCATCAGACGACGCG 360
Qy      136 CysLeuSerIleLeuTyrProIleTyrPheIleCysArgArgProArgTyrLeuSerSer 155
Db      361 TGCCGTGCTCATCTCTGAGCCCATCTGTAACCATGCGCCGCGCCCAATACCTTCATCG 420
Qy      156 ValMetCysValLeuLeuTyrPalaLeuSerLeuLeuArgSerIleLeuGluTyrMetPhe 175
Db      421 GTCATGTGTGTCCTGCTGCGGCGCTGTCCCTGCTGCGGAGTATCCTGAGTGTGATGTC 480
Qy      176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThr 195
Db      481 TGTGACTTCCCTGTTAGTGTGCTGATTCGTTGTGTGTAACGTCAGATTTCATTACA 540
Qy      196 ILeAlaThrLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuVal 215
Db      541 ATGCGCTGCGCTGTTTTTTTATATGTGTGTCTCTGTGTGTGTGTGTGTGTGTGTGTGT 600
Qy      216 ArgIleLeuCysGlySerArgIleGlyMetProLeuThrArgLeuTyrValThrIleLeu 235
Db      601 AGGATTTCTGTGTGATCCCGAGATGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
Qy      236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGluTyrPalaLeuPhe 255
Db      661 ACAGTGTGTGTCTTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Qy      256 ArgIleHisLeuAspTyrPlyValLeuPheCysHisValHisLeuValSerIlePheLeu 275
Db      721 AGGATTCACCTGATGTGAAAGTCTTATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Qy      276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGlu 295
Db      781 TCCGCTCTTAAACAGCAGTGTGCAACCCCATCATTTCTTTCGTGTGTGTGTGTGTGTGT 840
Qy      296 ArgGlnAsnArgGlnAsnLeuValLeuValLeuGlnArgAlaLeuGlnAsnThrProGlu 315
Db      841 CGTCAAATATGAGCAAGACCTGAGCTGTCTCCAGAGGCTGTGTGTGTGTGTGTGTGTGT 900
Qy      316 ValAspGluGlyGlyTyrPheProGlnGluThrLeuGluLeuSerGlySerArgLeu 335
Db      901 GTGATGTAGAGGTGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
Qy      336 GluGln 337
Db      961 GAGCAG 966

```

RESULT 9
ABT04875
ID ABT04875 standard; cDNA; 969 BP.
XX
XX ABT04875;
XX

DT 11-OCT-2002 (first entry)
XX
XX Human G protein coupled receptor hrup37 coding sequence.
XX
XX Human; G-protein coupled receptor; GPCR; hrup28; hrup29; hrup30; hrup31;
KM hrup32; hrup33; hrup34; hrup35; hrup36; hrup37; gene; ss.
XX Homo sapiens.
XX
XX MO200242461-A2.
XX
XX 30-MAY-2002.
PD

```

XX      26-NOV-2001; 2001KW-US044386.
PE      27-NOV-2000; 2000US-0253404P.
XX      PR 12-DEC-2000; 2000US-025346P.
XX      PR 20-FEB-2001; 2001US-027026P.
XX      PR 20-FEB-2001; 2001US-027028P.
XX      PR 06-APR-2001; 2001US-0282032P.
XX      PR 06-APR-2001; 2001US-0282356P.
XX      PR 06-APR-2001; 2001US-0282358P.
XX      PR 14-MAY-2001; 2001US-0290917P.
XX      PR 31-JUL-2001; 2001US-0309208P.
PA      (AREN-) ARENA PHARM INC.
PI      Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
XX      WPI; 2002-566565/60.
XX      DR P-PSDB; ABU04077.
XX      PT Novel endogenous and non-endogenous versions of G protein-coupled
XX      PT receptor useful for identification of candidate compounds as receptor
XX      PT agonists or antagonists for use as therapeutic agents.
XX      PS Claim 39; Page 74; 84pp; English.
XX      CC The present invention provides the protein and coding sequences of
XX      CC several human G-protein coupled receptors (GPCRs). These can be used in
XX      CC the identification of candidate compounds as receptor agonists or inverse
XX      CC agonists having applicability as therapeutic agents. The present sequence
XX      CC is a GPCR coding sequence of the invention
XX      SQ Sequence 969 BP; 171 A; 280 C; 250 G; 268 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8,34e-149 Length: 969
Score: 1688.00 Matches: 322
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.75% Indels: 0
DB: 6 Gaps: 0

US-09-867-570-2 (1-337) x ABT04875 (1-969)
Qy      16 MetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGlyArgGlu 35
Db      1 ATGATTCACACCATCCCAAGTCTTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60
Qy      36 ThrProCysTyrIleGlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
Db      61 ACTCTGTGTAAACAGCAGACCTGAGCTTCAACGCGGTGTGTGTGTGTGTGTGTGTGTGTGT 120
Qy      56 AlaLeuThrGlyAsnAlaValValLeuThrLeuGluGlyCysArgMetArgAsnAla 75
Db      121 GCGGTGACAGAAACGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180
Qy      76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95
Db      181 GTCTCCATCTACATCTCACTGCTGCGCCGCACTTCTCTTCTTACGCGCACTT 240
Qy      96  ILeCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
Db      241 ATATGTTGCGCGGTACCGCTCATCATATATCGGCATCTCCAAATCTCAGTCC 300
Qy      116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
Db      301 GTATGACCTTTCCCTTACTTATAGGCTTAAGCATGCTGACGCGCATCAGACGACGCG 360
Qy      136 CysLeuSerIleLeuTyrProIleTyrPheIleCysArgArgProArgTyrLeuSerSer 155
Db      361 TGCCGTGCTCATCTCTGAGCCCATCTGTAACCATGCGCCGCGCCCAATACCTTCATCG 420

```

QY 156 ValMetCysValIleuLeuThrAlaIleuSerIleuLeuArgSerIleuLeuGluIleuThrPheMetPhe 175
 DB 421 GTCATGTCGTCCTCTGAGCCCTGTCCTGCTGCGGAGATCTCGAGTGGATGTC 480
 QY 176 CysAspPheLeuPheSerGlyAlaAspSerValIleuPheGluIleuThrSerAspPheIleuThr 195
 DB 481 TGTGATCTTCCTGTTAGTGTCTGATGTTCTGTTGGTGTGAAACGTCAATTTCAATTA 540
 QY 196 IleAlaTrpLeuValPheLeuCysValValIleuCysGlySerSerIleuValIleuVal 215
 DB 541 ATCCGGTGGCTGGTGTGTTTATGTTGTTCTCTGTCGATCCAGCCGTGCTCTGCTG 600
 QY 216 ArgIleuLeuCysGlySerArgIleuMetProIleuThrArgIleuThrValThrIleuLeu 235
 DB 601 AGGATTCCTGTGGATCCCGGAGATGCCGTCGACCAAGCTGTACCTGACCATCTCTTC 660
 QY 236 ThrValIleuValPheLeuLeuCysGlyLeuProPheGlyIleuIleuThrAlaLeuPheSer 255
 DB 661 ACAGTGTGCTGCTCTCTCTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 256 ArgIleuHisLeuAspTrpIleuValIleuPheCysHisValHisLeuValSerIleuLeu 275
 DB 721 AGGATCCACCTGATGTAAGCTTATTTGTCATGTCATCTAGTTCATTTCCATTTTCTG 780
 QY 276 SerAlaLeuAsnSerSerAlaAsnProIleuIleuThrPhePheValGlySerPheArgGln 295
 DB 781 TCCGCTCTTACAGAGTCGACCCCAATTCATTTCTCTGTCGCTGCTCTTCTTCTGCTG 840
 QY 296 ArgGlnAsnArgGlnAsnLeuLeuValIleuGlnArgAlaLeuGlnAsnTrpProGln 315
 DB 841 CGTCAAAATAGGCAGAACCTGAAAGCTGTTCTTCAGAGGCTCTGACAGACAGCCCTGAG 900
 QY 316 ValAspGluGlyGlyGlyTrpLeuProGlnGluIleuThrLeuGluLeuSerGlySerArgGln 335
 DB 901 GTGGATGGAAGGTGAGGGTGGCTTCTCTCAGGAAACCTGAGCTGTGCGGAGACAGATTG 960
 QY 336 GluGln 337
 DB 961 GAGCAG 966
 Db
 RESULT 10
 ADO30080
 ID ADO30080 standard; cDNA; 969 BP.
 AC ADO30080;
 XX 29-JUL-2004 (first entry)
 DT
 DE Human GPCR MRGK3 polynucleotide, SEQ ID NO:1182.
 XX
 KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KW transgenic mouse; neurological disorder; adrenal gland disorder;
 KW colon disorder; intestinal disorder; cardiovascular disorder;
 KW muscular disorder; blood disorder; immune disorder; bone disorder;
 KW joint disorder; metabolic disorder; nutritive disorder; cancer;
 KW kidney disorder; liver disorder; lung disorder; breast disorder;
 KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KW thymus disorder; thyroid disorder; antiparkinsonian; antianemic;
 KW cytosolic; antiinflammation; vasorelaxant; antidiabetic;
 KW CNS; central nervous system; respiratory; antidiarrhoeal; antidiabetic;
 KW vitruicide; hepatotropic; antibacterial; antianemic; antiseborrhoeic;
 KW dermatological; antitumor; antichyroid; antiallergic; anorectic;
 KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2004040000-A2.
 XX
 PD 13-MAY-2004.
 XX
 PF 09-SEP-2003; 2003WO-US028226.

XX 09-SEP-2002; 2002US-0409303P.
 PR 09-APR-2003; 2003US-0461329P.
 PA (PRIM-) PRIMAL INC.
 PI Galanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F,
 PI Madisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;
 XX WPI, 2004-390329/36.
 DR P-PSDB; ADO29705.
 PT Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.
 PS Claim 151; SEQ ID NO 1182; 542pp; English.
 XX
 XX The invention relates to human and mouse G protein-coupled receptors
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention; methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice; kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridize to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of
 CC diseases including neurological disorders (e.g., Alzheimer's disease,
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders; blood disorders (e.g.,
 CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
 CC nucleic acid of the invention. Note: The full sequence data for this
 CC patent did not form part of the printed specification, those sequences
 CC not shown were obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 969 BP; 171 A; 280 C; 250 G; 268 T; 0 U; 0 Other;
 SQ
 Alignment Scores:
 Pred. No.: 8,34e-149 Length: 969
 Score: 1688.00 Matches: 332
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 95.75% Indels: 0
 DB: 12 Gaps: 0
 US-09-867-570-2 (1-337) X ADO30080 (1-969)
 QY 16 MetAspSerThrIleProValIleuGlyTrpIleuThrProIleuAsnGlyValArgGluGln 35
 DB 1 ATGATTTCAACCAATCCAGCTTGGGTACAGAACTGACCAATCAACGAGCTGAGGAG 60
 QY 36 ThrProCysTrpArgGlnThrLeuSerPheTrpGlyLeuThrCysIleValSerLeuVal 55
 DB 61 ACTCCTTGCTACAGAGAACCTGAGCTTCAAGGGGCTGACGTGATCTTCCCTTGTGC 120
 QY 56 AlaLeuThrGlyValAsnIleValIleuTrpLeuLeuGlyCysArgMetArgAsnAla 75

QY 156 ValMetCysValLeuLeuThrAlaLeuSerLeuLeuArgSerIleLeuGluThrPhe 175
 Db 421 GTCATGTGTCTGCTCTGGGCCCCCTGTCCTGCGGAGATCTCGAGTGAAGTTC 480
 QY 176 CysAspPheLeuPheSerGlyAlaAspSerValITP-CysGluThrSerAspPheIleThr 195
 Db 481 TGTGACTTCTGCTTGTAGTGTCTGATTTCTGTTGGTGTGAACGTCAGATTTCATTACA 540
 QY 196 IleAlaTrpLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuVal 215
 Db 541 ATCCGGCTGGCTGTTTATATGTGTGCTCTGCGGCTCCAGCTGCTCTGCTGCTGTC 600
 QY 216 ArgIleLeuCysGlySerArgGlyMetProLeuThrArgLeuThrValThrIleLeuLeu 235
 Db 601 AGGATTCCTGTGTGATCCCGGAAGATGCGCTGACCGAGCTGTACGTGACCATCTCTC 660
 QY 236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGluThrAlaLeuPheSer 255
 Db 661 ACAGTGTGCTGCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 256 ArgIleHisLeuAspTrpIleValLeuPheCysHisValHisLeuValSerIlePheLeu 275
 Db 721 AGGATCCACCTGGATTGGAAAGCTTATTTGTGCATGTCATCTAGTTCCATTTCCTG 780
 QY 276 SerAlaLeuAspSerSerAlaAspProIleIleThrPhePheValGlySerPheArgGlu 295
 Db 781 TCCGCTCTTAACAGAGTGCACCCCATCATTTACTTCTCTGCGGCTCTCTTTAGGCGAG 840
 QY 296 ArgGluAsnArgGluAsnLeuLysLeuValLeuGluAlaArgAlaLeuGluAspThrProGlu 315
 Db 841 CGTCAAAATAGCGCAACACTGAAAGCTGCTTCTCCAGAGGCTCTGACGACAGCCCTGAG 900
 QY 316 ValAspGluGlyGlyIleTrpLeuProGluGluThrLeuGluLeuSerGlySerArgLeu 335
 Db 901 GTGGATGAAAGGTGAGGCTGCTCTCTCAGAGAAACCTGAGCTGTGCGGAGACAGATTG 960
 QY 336 GluGlu 337
 Db 961 GAGCAG 966

RESULT 12
 ADO44605
 ID ADO44605 standard; DNA; 966 BP.
 XX
 AC ADO44605;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Human HIT213 protein encoding DNA.
 XX
 KM HIT213; transgenic; G protein-coupled receptor; GPCR; ophthalmological;
 KW cytoskeletal; nephrotoxic; antiinflammatory; dermatological; analgesic;
 XX vulnary; neuroprotective; human; gene; ds.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..966
 FT /*tag= a
 FT /product= "HIT213"
 XX
 XX WO2004039972-A1.
 XX
 PD 13-MAY-2004.
 XX
 PF 28-OCT-2003; 2003WO-JP013781.
 XX
 PR 29-OCT-2002; 2002JP-00314141.
 XX
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Katsuo Y, Watanabe T, Yasuhara Y, Mori I, Takeomi S;

XX
 DR WPI: 2004-376191/35.
 DR P-PSDB; ADO44604.
 XX
 PT HIT213 protein, encoded DNA and transgenic animals for clarifying
 PT pathological mechanism, developing therapeutic methods and screening
 PT preventives or remedies for related diseases e.g. cataract, cancer, and
 PT dermatitis.
 XX
 XX disclosure; SEQ ID NO 4; 161bp; Japanese.
 XX
 CC The invention relates to a non-human mammal that carries a DNA integrated
 CC with a foreign HIT213 or its mutant gene, or a part of it. The non-human
 CC animal is particularly a rat. Such gene shows phenotypes of e.g. cataract
 CC onset, transient skin rash and proliferation-promoting activity. The
 CC foreign HIT213 gene is a gene that encodes a G protein-coupled receptor
 CC (GPCR) protein HIT213. The protein, its encoded DNA and constructed
 CC transgenic animals are useful for clarifying pathological mechanism,
 CC developing therapeutic methods and screening preventives or remedies for
 CC related diseases e.g. cataract, cancer, and dermatitis. The present
 CC sequence represents a DNA encoding a human HIT213 protein.
 XX
 SQ Sequence 966 BP; 171 A; 280 C; 248 G; 267 T; 0 U; 0 Other;
 XX

Alignment Scores:
 Pred. No.: 2,45e-148 Length: 966
 Score: 1683.00 Matches: 321
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.69% Mismatches: 0
 Query Match: 95.46% Indels: 0
 DB: 12 Gaps: 0

US-09-867-570-2 (1-337) x ADO44605 (1-966)

QY 16 MetAspSerThrIleProValLeuGlyThrGluLeuThrProIleAsnGlyArgGlu 35
 Db 1 ATGGATTCAACCATCCAGCTTGGGTACAGAACCAACCAATCAACGAGCTGAGGAG 60
 QY 36 ThrProCysTrpIleGluThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
 Db 61 ACTCCTGTACAGAGACCCCTGAGCTTACAGGAGCTGACGTCGATGATTCCTTGTTC 120
 QY 56 AlaLeuThrGlyAsnAlaValValLeuTrpLeuLeuGlyCysArgMetArgArgAsnAla 75
 Db 121 GCGCTACAGAAACCGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 QY 76 ValSerIleTrpIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95
 Db 181 GTCTCCATCTACATCTCAACCTGGTGGCGGCGGCACTTCTCTTCAAGCGCCACAT 240
 QY 96 IleCysSerProLeuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
 Db 241 AATGTTCGCCGTATGCGCTCAATATATCCCAATCCCATCTCCAAATCTCACTGCT 300
 QY 116 ValMetThrPheProTrpPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
 Db 301 GTGATACCTTTCCTTCCATCTTATAGCCCTTAACAGATGTGAGCCATCAGACGAGGCC 360
 QY 136 CysLeuSerIleLeuTrpProIleTrpIleGlyCysArgArgProArgTrpLeuSerSer 155
 Db 361 TGGCTGTCACTCTGCGCCCATCTGTGTACCTGCGCCGCCCAAGATACCTGTCACTG 420
 QY 156 ValMetCysValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGluThrPhePhe 175
 Db 421 GTCATGTGTCTGCTCTGGGCCCCCTGTCCTGCGGAGATCTCGAGTGAAGTTC 480
 QY 176 CysAspPheLeuPheSerGlyAlaAspSerValITP-CysGluThrSerAspPheIleThr 195
 Db 481 TGTGACTTCTGCTTGTAGTGTCTGATTTCTGTTGGTGTGAACGTCAGATTTCATTACA 540
 QY 196 IleAlaTrpLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuVal 215
 Db 541 ATCCGGCTGGCTGTTTATATGTGTGCTCTGCGGCTCCAGCTGCTCTGCTGCTGTC 600

QY 216 Arg11leuCyseGlySerArglyMetProleuThraArgleuTyVal1Thr1leuLeu 235
 Db 601 AGGATTCCTGTGGATCCCGAAGATCCGCTGACCGAGGTGTACGTGACCATCTCTC 660
 QY 236 ThrValleuValPheLeuLeuCyseGlyLeuProphlegly1legIntPalaLeuPheSer 255
 Db 661 ACAGTGTGATCTCTCTCTCTGAGGCTGCTTGGCATTCAGTGGCCCTCTTTTC 720
 QY 256 Arg11leuLeuAspTrpLyValLeuPheCyseHisValHisLeuValSer1lePheLeu 275
 Db 721 AGGATTCACCTGGATGGAAAGCTTATTTTGTATGTGATGATCACTAGTTTCATTTCTG 780
 QY 276 Ser11leuLeuAsnSerSer11aAsnPro11e11eTyPhePheVal1GlySerPheArgLeu 295
 Db 781 TCGGCTCTTAAACAGCAGTGGCCAAACCCATCATTTACTTCTTCTGAGGCTCTTTAGGCAG 840
 QY 296 ArgGlnAsnArgGlnAsnLeuValLeuGln1aLeuGln1aAspThrProGlu 315
 Db 841 CGTCAAAATAGGCAACCTGAAAGTGTCTTCCAGAGGGCTTGCAGAGACCGCTGAG 900
 QY 316 ValAspGln1GlyGlyTyTrpLeuProGlnGluThrLeuGlnLeuSerGlySerArgLeu 335
 Db 901 GTGGATGAAGTGGAGGGTGGCTTCTTCCAGAAACCTGAGAGCTGTGGAGAGCATTTG 960
 QY 336 GlnGln 337
 Db 961 GAGCAG 966

RESULT 13

ID AAA29812
 AAA29812 standard; cDNA; 969 BP.

AC AAA29812;

DT 18-AUG-2000 (first entry)

DE Human G protein-coupled receptor hMT7213V encoding cDNA SEQ ID NO:4.

KW Human; G protein-coupled receptor; hippocampus; diagnosis; screening;

KW genetic disease; cellular function regulation; ss.

OS Homo sapiens.

EH Key Location/Qualifiers

FT CDS 1..969

FT /*tag= a

FT /product= "G protein-coupled receptor"

PN WO200020455-A1.

PD 13-APR-2000.

PF 30-SEP-1999; 99WO-JP005366.

PR 01-OCT-1998; 98UP-00279535.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Watanabe T, Terao Y, Matsui H;

DR WPI, 2000-303747/26.

DR P-PSDB; AAY90762.

PT Human-derived G protein-coupled protein and encoding nucleic acid, useful

PT e.g. in determining ligands and treatment of diseases associated with

PT dysfunction of the protein.

PS Claim 6; Page 94-95; 97pp; Japanese.

CC The present sequence encodes a human-derived G protein-coupled protein

CC designated hMT7213V, which is isolated from the human hippocampus. The G

CC protein-coupled receptor can be used for preventing, treating and

CC diagnosing genetic diseases associated with G protein-coupled protein,
 CC and for regulating cellular functions. The protein can be used to prevent
 CC and treat disorders associated with G protein-coupled protein gene
 CC dysfunction. It can also be used to identify G protein-coupled protein
 CC ligands and generating antibodies and antisera against the protein. It is
 CC also useful in constructing recombinant receptor protein expression
 CC systems, developing receptor-binding assay systems and screening drug
 CC candidates, and can be used as a probe in the genetic diagnosis of G
 CC protein-coupled protein disorders

SO Sequence 969 BP; 172 A; 280 C; 249 G; 268 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,46e-148	Length:	969
Score:	1683.00	Matches:	321
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.69%	Mismatches:	0
Query Match:	95.46%	Indels:	0
DB:	3	Gaps:	0

US-09-867-570-2 (1-337) x AAA29812 (1-969)

QY 16 MetAspSerThr11eProVal1leuGlyThrGluLeuThrPro11eAsnGlyArgGlnGlu 35
 Db 1 ATGGATTCACCACTCCAGTCTTGGTACAGAACTTGACCAATCAACGAGCGAGAG 60
 QY 36 ThrProCyseTyTrpLyGlnThrLeuSerPheThrGlyLeuThrCyse1leValSerLeuVal 55
 Db 61 ACTCCTTGCTAACAGACGACCGTACGCTTACCGGGGCTGACGTGCATCGTTCCCTTGTG 120
 QY 56 AlaLeuThrGlyAsnAlaValLeuThrLeuGlnGlyCysArgMetCysArgAsnAla 75
 Db 121 GCGCTGACAGAAACGGGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 76 ValSer11eTyTrp11eLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHis11e 95
 Db 181 GTCTTCATCTTACATCTTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 QY 96 11eCyseSerProleuArgLeu1eAsn1eArgHisPro11eSer11eLeuSerPro 115
 Db 241 ATATGTTGGCGGTAGGCTCATCAATATCGGCATTCATCAATATCTCATGTCCT 300
 QY 116 ValMetThrPheProTyTrpPhe11eGlyLeuSerMetLeuSerAla11eSerThrGluArg 135
 Db 301 GTGATGACCTTCCCTTCTTATAGGCTTAAAGCATGTCTGAGCGGCATCAGACCGAGCGC 360
 QY 136 CysLeuSer11eLeuTrpPro11eTrpTyHisCysArgArgProArgTyTrpLeuSerSer 155
 Db 361 TGCTGTCAATCTGTGGCCATCTGTGACACTGCGCGCCGCCAAGATACCTGTCAATCG 420
 QY 156 ValMetCysVal1leuLeuTrpAlaLeuSerLeuLeuArgSer11eLeuGluTrpMetPhe 175
 Db 421 GTCATGTGTCTCGTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 176 CysAspPheLeuPheSerGlyAlaAspSerValTrpCyseGluThrSerAspPhe11eThr 195
 Db 481 TGTGACTTCTCTGTTAGTGTGCTAATCTGTGTGGTGTGTAACGTGAGATTTTCATTACA 540
 QY 196 11eAlaTrpLeuValPheLeuCyseVal1leuGlySerSer11eLeuVal1leuVal 215
 Db 541 ATGCGTGGCTGTTTATATGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 QY 216 Arg11leuCyseGlySerArglyMetProleuThraArgleuTyVal1Thr1leuLeu 235
 Db 601 AGGATTCCTGTGGATCCCGAAGATCCGCTGACCGAGGTGTACGTGACCATCTCTCTC 660
 QY 236 ThrValleuValPheLeuLeuCyseGlyLeuProphlegly1legIntPalaLeuPheSer 255
 Db 661 ACAGTGTGATCTCTCTCTCTGAGGCTGCTTGGCATTCAGTGGCCCTCTTTTC 720
 QY 256 Arg11leuLeuAspTrpLyValLeuPheCyseHisValHisLeuValSer1lePheLeu 275
 Db 721 AGGATTCACCTGGATGGAAAGCTTATTTTGTATGTGATGATCACTAGTTTCATTTCTG 780

QY 276 SerAlaLeuAenSerSerAlaAnProIleIleTyrPhePheValGlySerPheArgGln 295
 Db 781 TCCCTCTTAAcAGAGGTCACCAcCCCAcATTACTTCTTGCGGCTCTTAAAGGcAG 840
 QY 296 ArgGlnAenArgGlnAenLeuValLeuGlnArgAlaLeuGlnAenPheThrProGln 315
 Db 841 CGTCAAAATAGGcAGACCTGAGcCTGCTTCCAGAGGcCTCTGAGAGACACGCTTgAG 900
 QY 316 ValAspGlnGlyGlyTyrIleuProGlnGlnIuThrLeuGlnLeuSerGlySerArgLeu 335
 Db 901 GTGATGAAGGTGAGGcGTGCTCTCAGAAcCCTGAGcCTGCGGAAAGcAGATTg 960
 QY 336 GlnGln 337
 Db 961 GAGcAG 966

RESULT 14
 ADF29104
 ID ADF29104 standard; cDNA, 969 BP.
 AC ADF29104;
 XX 12-FEB-2004 (first entry)
 DT 12-FEB-2004 (first entry)
 XX
 DE Human GPCR R-14-encoding cDNA.
 XX Human, R-14; GPCR; G protein coupled receptor;
 KW trabecular meshwork tissue; ocular outflow; antagonist;
 KM ocular hypotensive; drug screening; elevated intraocular pressure;
 KW glaucoma; ophthalmological; gene; ss.
 XX Homo sapiens.
 OS
 PN W02003080659-A1.
 XX
 PD 02-OCT-2003.
 XX
 PF 27-MAR-2003; 2003WO-CA000444.
 XX
 PR 27-MAR-2002; 2002US-0367513P.
 XX
 PA (THER-) THERATECHNOLOGIES INC.
 XX
 PI Peri KG, Moffett S, Adrian D;
 XX
 DR WPI; 2004-053019/05.
 DR P-PSDB; ADF29105.
 XX
 PT New substantially pure R-14 polypeptide, useful as drug targets for
 PT lowering intraocular pressure and for treating condition such as
 PT glaucoma.
 XX
 PS Claim 22; SEQ ID NO 1; 85bp; English.
 XX
 CC The invention relates to a human GPCR (G protein coupled receptor)
 CC designated R-14 (ADF29105), nucleic acids encoding it (ADF29104), and R-
 CC 14 peptide antagonists (ADF29106-ADF29108). The invention also
 CC encompasses vectors and host cells comprising R-14 nucleic acids, and
 CC methods of screening for R-14 antagonists. The R-14 receptor is expressed
 CC in trabecular meshwork tissue, and is associated with a role in ocular
 CC outflow. Inhibition of the receptor results in a reduction of basal
 CC intraocular pressure, making the R-14 receptor a useful target for
 CC screening for ocular hypotensive drugs. The R-14 peptide antagonists are
 CC useful for reducing intraocular pressure for the treatment of conditions
 CC associated with elevated intraocular pressure such as glaucoma and
 CC related conditions. The R-14 receptor, and host cells expressing an R-14
 CC polynucleotide, may be used in screening for R-14 receptor antagonists.
 CC The present sequence is related to the invention.
 XX
 SQ Sequence 969 BP; 171 A; 280 C; 250 G; 268 T; 0 U; 0 Other;

Alignment Scores:

Prod. No.: 2,466-148 Length: 969
 Score: 1683.00 Matches: 321
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.69% Mismatches: 0
 Query Match: 95.46% Indels: 0
 DB: 12 Gaps: 0

US-09-867-570-2 (1-337) x ADF29104 (1-969)

QY 16 MetAspSerThrIleProValLeuGlyThrGlnLeuThrProIleAenGlyArgGln 35
 Db 1 ATGGAATTCAACCATCCAGTCTTGcGTACAGAACTGACCAATCAAGAGcGTGAGAG 60
 QY 36 ThrProCysTyrIleGlnThrIleuSerPheThrGlyLeuThrCysIleValSerLeuVal 55
 Db 61 ACTCTTGCTACAGAGACcCTGAGcCTTACGAGGcCTGACGTGcATGCTTCTTCTGTC 120
 QY 56 AlaLeuThrGlyAsnAlaValValLeuThrPLeuLeuGlyCysArgMetArgAsnAla 75
 Db 121 GCGGTGACAGAGAGcCGGcTGTGCTGcGTGCTGcCGcCTGcCGcATGCGAGAAcGCT 180
 QY 76 ValSerIleTyrIleLeuAenLeuValAlaAlaAspPheLeuSerGlyValIle 95
 Db 181 GTCTCCATCTACATCTCTCAACCTGcGTGcCGcGcGcCTTCTTCTTACGCGcCACT 240
 QY 96 IleCysSerProLeuArgLeuIleAenIleArgHisProIleSerIleLeuSerPro 115
 Db 241 ATATGTTCCCGCTTACGCTTCAATCAATTCGCCATTCcATTCcAAATCTCTCACTCT 300
 QY 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGlnArg 135
 Db 301 GTGATGACCTTCCcATCTTATAGcCGCTAAGcATGCTGAGcGCATCAGACcGAGcGC 360
 QY 136 CysLeuSerIleLeuThrProIleTyrIleHisCysArgArgProArgTyrLeuSerSer 155
 Db 361 TGCTGTCTCATCTGcGTGcCGcCATCTGTAACACTGcCGcCGcCGcCGcCATCTGCTG 420
 QY 156 ValMetCysValLeuLeuThrPalaLeuSerLeuLeuArgSerIleLeuGlnTyrPhePhe 175
 Db 421 GTCATGTGTGCTGCTGCTGcCGcCGcCTGCTGcCGcAGATTCCTGAGTGAATGTC 480
 QY 176 CysAspPheLeuPheSerGlyAlaAspSerValTyrCysGlnThrSerAspPheIleThr 195
 Db 481 TGTGACTTCTGTTATAGcGTGCTGATTCGTGTTGcGTGAACcGTGAGATTCTTACA 540
 QY 196 IleAlaTyrPLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuVal 215
 Db 541 ATCGcGTGCTGcGTGTTTATGTGTGcGTCTGcGTGcGTGcGTGcGTGcGTGcGTG 600
 QY 216 ArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuTyrValThrIleLeuLeu 235
 Db 601 AGGATTCCTGTGAGATCCGAGAGATGcCGcGTGACcAGcCTTACGTACACCTCTCTC 660
 QY 236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnThrPalaLeuPheSer 255
 Db 661 AAGAGCTGCTGTCTCTCTCTGcGTGcCGcCTTGGcATTCAGAGGcCGcCTGTTTCC 720
 QY 256 ArgIleHisLeuAspTyrIleValLeuPheCysHisValHisLeuValSerIlePheLeu 275
 Db 721 AGGATTCACCTGATGAGATGTAATCTTATTTGcATGTGcCATCTGATCTGATTTCCG 780
 QY 276 SerAlaLeuAenSerSerAlaAnProIleIleTyrPhePheValGlySerPheArgGln 295
 Db 781 TCCGCTCTTAAcAGAGGTCACCAcCCCAcATTACTTCTTGCGGCTCTTAAAGGcAG 840
 QY 296 ArgGlnAenArgGlnAenLeuValLeuGlnArgAlaLeuGlnAenPheThrProGln 315
 Db 841 CGTCAAAATAGGcAGAAcCTGAGcCTGCTTCCAGAGGcCTCTGAGAGACACGCTTgAG 900
 QY 316 ValAspGlnGlyGlyTyrIleuProGlnGlnIuThrLeuGlnLeuSerGlySerArgLeu 335
 Db 901 GTGATGAAGGTGAGGcGTGCTCTCAGAAcCCTGAGcCTGCGGAAAGcAGATTg 960

QY 336 GluGln 337
 Db 961 GAGCAG 966

RESULT 15
 ID AAA70342 standard; cDNA; 969 BP.
 XX AAA70342;
 AC AAA70342;
 XX 19-DEC-2000 (first entry)
 DE Human novel G-protein coupled receptor #1 coding sequence.
 XX Human; novel G-protein coupled receptor; signal transduction;
 KM disease diagnosis; drug screening; disease therapy; ss.
 XX Homo sapiens.
 OS
 PH Key Location/Qualifiers
 FT CDS 1..969
 FT /*tag= a
 FT /product= "novel G-protein coupled receptor #1"

PN M0200040724-A1.
 PD 13-JUL-2000.
 XX 04-JAN-2000; 2000MO-US000052.
 PF 04-JAN-1999; 99US-0114666P.
 PR 14-JAN-1999; 99US-0115828P.
 XX (LEXI-) LEXICON GENETICS INC.
 XX
 PI Nehls M, Wattler F;
 XX WPI; 2000-465986/40.
 DR P-PSDB; AABI4846.
 XX
 PT New polynucleotides encoding novel G-protein coupled receptors useful for
 PT diagnosis, drug screening, clinical trial monitoring and for the
 PT treatment of physiological or behavioral disorders.
 XX
 XX Claim 1; Page 52-53; 61pp; English.
 XX
 XX The present sequence is the coding sequence for a novel human G-protein
 CC coupled receptor (NGPCR). These proteins are involved in signal
 CC transduction pathways in many cases. The protein contains seven
 CC transmembrane domains, and is expressed in human testis, mammary gland
 CC and salivary gland tissue. The gene, its protein, agonists, antagonists
 CC and antibodies can be used to diagnose and treat diseases associated with
 CC the inappropriate expression or expression of mutant versions of the
 CC protein, for screening for drugs which can be used in the same manner,
 CC and for elucidating the function of the protein

SQ Sequence 969 BP; 172 A; 282 C; 249 G; 266 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No: 1,866-146 Length: 969
 Score: 1663.00 Matches: 320
 Percent Similarity: 99.38% Conservative: 0
 Best Local Similarity: 99.38% Mismatches: 2
 Query Match: 94.33% Indels: 0
 DB: 3 Gaps: 0

US-09-867-570-2 (1-337) x AAA70342 (1-969)

QY 16 MetAspSerThrIleProValIleuGlyThrGluLeuThrProIleAsnGlyArgGluGln 35
 Db 1 ATGGATTCAACCATCCAGCTTGGGTACAGAACTGACCAATCAACGACGAGAGAG 60

QY 36 ThrProCysTyrIysGlnThrLeuSerPheThrGlyLeuThrCysIleValSerLeuVal 55

Db 61 ACTCTTGCTTACAGACACCTGAGCTTACAGGGGCTGACGTCATGTTCCCTTGTC 120
 QY 56 AlaLeuThrGlyAsnAlaValIleuTrpLeuLeuGlyCysArgMetArgArgAsnAla 75
 Db 121 GCGCTGACAGAAACCGGTTGTGCTCTGAGCTCCTGGGCTGCCGACATCGCAGAAAGCT 180
 QY 76 ValSerIleTyrIleLeuAsnLeuValAlaAlaAspPheLeuPheLeuSerGlyHisIle 95
 Db 181 GTCTCCATCTTACATCCCAACCTGGTCGCGCGGACCTTCTCTTCTTACCGGCACTT 240
 QY 96 IleCysSerProIleuArgLeuIleAsnIleArgHisProIleSerIleLeuSerPro 115
 Db 241 ATACGTTCGCGGTACGCTCATCATATATCCGCAATCCATCTCCAAATCTCATGTCCT 300
 QY 116 ValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArg 135
 Db 301 GTGATGACCTTCCCTTACTTATAGGCTTAAAGCATGCTGAGCGCATCAGCAGAGCGC 360
 QY 136 CysLeuSerIleLeuTrpProIleTyrIleCysArgArgProArgTyrLeuSerSer 155
 Db 361 TGCCTGTCCATCCTGTGGCCCATCTGGTACACATGCGCCGCCCAATACCTGTCAATCG 420
 QY 156 ValMetCysValIleuLeuTrpAlaLeuSerLeuLeuArgSerIleuGluTrpMetPhe 175
 Db 421 GTCATGTGTCTCTGCTCTGAGCCCTGCTGCTGCGAGATATCTTGAGTGAATGTC 480
 QY 176 CysAspPheLeuPheSerGlyAlaAspSerValTrpCysGluThrSerAspPheIleThr 195
 Db 481 TGTGACTTCTCTGTTTATGTTGCTGATGCTGCTTGGTGAAGCTGAAATTCATTACA 540
 QY 196 IleAlaTrpLeuValPheLeuCysValValLeuCysGlySerSerLeuValLeuVal 215
 Db 541 ATCGCGTGTCTGTTTATGTTATGTTGTTCTCTGTGGGTCCAGCTGTGCTGCTGTC 600
 QY 216 ArgIleLeuCysGlySerArgIysMetProLeuThrArgLeuTyrValIleLeuLeu 235
 Db 601 AGGATCTCTGTGATATCCGAGAGATGCCGTGACAGGCTGATGATGATCATCTCTC 660
 QY 236 ThrValLeuValPheLeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSer 255
 Db 661 ACGATGCTGCTTCTCTCTCTGAGGCTGCCCTTGGCATTCAGTGGGCTGTTTCC 720
 QY 256 ArgIleHisLeuAspTrpIysValLeuPheCysHisValHisLeuValSerIlePheLeu 275
 Db 721 AGGATCCACCTGGAATTGAAAGTCTTATTTGTCATGTCATGATCAATTCATTTCTC 780
 QY 276 SerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGlySerPheArgGln 295
 Db 781 TCCGCTCTTAAACAGAGTGCACACCCCATCATTTACTTTCGTGGGCTCCCTTAAAGCAG 840
 QY 296 ArgGlnAsnArgGlnAsnLeuValLeuValIleuGlnArgAlaLeuGlnAspThrProGlu 315
 Db 841 CGTCAAAATAGGCAAGAACTGAAAGCTGTTCTTCAAGAGGCTCTGCAGACACCCCTGAG 900
 QY 316 ValAspGluGlyGlyTyrTrpLeuProGlnGluThrLeuGluLeuSerGlySerArgLeu 335
 Db 901 GTGGATGAAGTGAAGGACAGCTTCCCTCAAGAAACCTCGAGGCTGTGCGAAGACGATTG 960

QY 336 GluGln 337
 Db 961 GAGCAG 966

Search completed: December 1, 2004, 23:05:59
 Job time: 505 secs

This Page Blank (uspto)

[illegible]

Email: sca1n@therys.com
 High quality sequence stop: 559.
 Location/Qualifiers
 1..764
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="HT1080"
 /clone_lib="Atherys RAGE library"
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN

Alignment Scores:

Pred. No.:	4.27e-95	Length:	764
Score:	1032.50	Matches:	205
Percent Similarity:	86.77%	Conservative:	18
Best local Similarity:	79.77%	Mismatches:	29
Query Match:	58.56%	Indels:	5
		Gaps:	1

US-09-867-570-2 (1-337) x BG198766 (1-764)

```

Oy      11etYrIleLeuAsnLeuValAlaAlaAspPheLeuSer-GlyHisIleIleCy 97
Db      764 ATCTACATCTCTCAACATGCGCGACAGACTTCTCTGCTCAGCGGCGCTTATATA 705
Oy      97  sSerProLeuArgLeuIleAsnIle-ArgHisProIleSerIysIleLeuSerProValM 117
Db      704 TTCCCTGTACTCTTCATCAGATGATCCCCCAACCATCTCTAAATCCCTTATCCGTGA 645
Oy      117 ectMrPheProTyPheHeIegIyLeuSerMetLeuSerAlaIleSerThrGluArgCyel 137
Db      644 TGATGTTTTCATCTTCAGCGCCCTGACGCTTTCTAAGTCCGTGACCGACCGAGCGCTGCC 585
Oy      137 eusSerIleLeuTrpProIleTrpIyHisCyAArgProArgIyLeuSerSerValM 157
Db      584 TGTCGCTCTGTGGGCCCATCTGTGTACCGGTGCCACCGCCCAACACACTGTACAGCGGTG 525
Oy      157 etCyValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleLeuGluTrpMetPheCysA 177
Db      524 TGTGTGTCCTGTCTGTGGGCCCTGTCTCGCTGCGAGCATCTCGAGATGTATGTG 465
Oy      177 sPheLeuPheSerGlyAlaAspSerValTrpCyGluThrSerApeHeIleThrIleA 197
Db      464 GCTTCCTGTTCAGTGTGCTGTGATTCGTGCTGTGTCMAACATCAGATTTCATCAGATCG 405
Oy      197 laTrpLeuValPheLeuCyValValLeuCyGlySerSerLeuValLeuLeuValArgI 217
Db      404 CGTGCGTAAATTTTTTATGTGTGTGTCTGTGGGTCCAGCCCGGCTCGTGTGATGAGA 345
Oy      217 leLeuCyGlySerArgIysMetProLeuThrArgLeuIyValThrIleLeuLeuThy 237
Db      344 TTCTCTGTGTGATCCCGAAGAACCGGTGACCAAGCGTGTACGTGACCATCTGTCTCAG 285
Oy      237 alLeuValPheLeuLeuCyGlyLeuProPheGlyIleGluTrpAlaLeuPheSerArgI 257
Db      284 TACTGTGTCTCTCTCTGTGTGGCCCTGTGGCATTCAGTTTTCCTTAATTTTATGGA 225
Oy      257 leHisLeuAspTrpIyValLeuPheCyHisValHisLeuValSerIlePheLeuSerA 277
Db      224 TTCACGTGACAGGGAAAGCTTATTTGTGACGTTCATCTAGTTCTTATTTCTGTGCCG 165
Oy      277 laLeuAsnSerSerAlaAsnProIleIleTyPhePheValGlySerPheArgGlnArg 297
Db      164 CTCTTAAAGCAGAGTCCCAACCCCATTTACTTCTTGCTGTGGGCTCTTTAGCAGCGTC 105
Oy      297 lAsnArgGlnAsnLeuIyLeuValLeuGlnArgAlaLeuGlnAspThrProGluValA 317
  
```

Db 104 AAAATAGGACAGAACTGAGTGTCTCCAGAGGGCTCTGACGAGCGGCTGAGGTG 45

Qy 317 spglluglvglyglttrpleuproglingluthrleugluleuSerGly 332

Db 44 ATGAGAGTGAGGAGGAGAAATCCTGAG-----GAAATGCTGGA 7

RESULT 4

BE439409 688 bp mRNA linear EST 25-JUL-2000

LOCUS HTML-025F1 HTML Homo sapiens cDNA, mRNA sequence.

DEFINITION BE439409

ACCESSION BE439409.1 GI:9438891

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS Gonzalez, P., Epstein, D.L. and Borras, T.

TITLE Characterization of gene expression in human trabecular meshwork using single-pass sequencing of 1060 clones

JOURNAL Invest. Ophthalmol. Vis. Sci. (2000) In press

COMMENT Contact: Pedro Gonzalez

Department of Ophthalmology

Duke University

Duke Eye Center, Erwin Rd, Box 3802, Durham, NC 27710-3802, USA

Tel: 919 681 4065

Fax: 919 684 8983

Email: pedro.gonzalez@duke.edu.

Location/Qualifiers

1..688

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="trabecular meshwork"

/clone_lib="HTM1"

ORIGIN

Alignment Scores:

Pred. No.: 3,42e-92 Length: 688

Score: 1003.50 Matches: 211

Percent Similarity: 96.80% Conservatve: 1

Best Local Similarity: 96.35% Mismatches: 7

Query Match: 56.92% Indels: 6

DB: 2 Gaps: 0

US-09-867-570-2 (1-337) x BE439409 (1-688)

Qy 8 ValleargleuglYpHeuSerMeCaSpSerThrlEProValleuglYThrgluleu 27

Db 36 GTCAACAGACTGGGTTTCTGAGCATGATTCACCATCCGACTTGGGTACAGAACTG 95

Qy 28 ThrProileanrglYArGluGluThrProCyTYrlyGlnThrlEusSerPheThrgly 47

Db 96 AACACCAATCAACGAGCGTGAAGAGACTCTTGCTTCAACAGACAGCCTGAGCTTCAAGGAG 155

Qy 48 LeuThrCySllleValSerleuValAlaLeuThrglyASnaAlaValleuThrpIleu 67

Db 156 CTGAGCGTGTGATGTTCCCTTGTGGCGTGAACGAAAGCGGTTGTCTGGCTCCCTG 215

Qy 68 GlyCyArGMeArGArGAsnaAlaValSerlleYrleleuASnleuValAlaAlaAp 87

Db 216 GGCTGCCGATCCGACGAAACCTGTCTCATTTACATCTCAACTGGTGGCGGCGAGC 275

Qy 88 PheleuPheleuSerGlyHisllelleCySserProleuThrgleuThleuThleuThleu 107

Db 276 TTCCTCTTCTTACGGGCGACCATTAATATGTTCCCGCTTACGCTCATCAATATCCGCAT 335

Qy 108 ProileSerleYrleleuSerProValMetThrPheProThrPheThleuThleuSerMe 127

Db 336 CCCATCTCCAAATCTCATGCTCTGTGATGACCTTTCCTACTTATAGGCTTAAGCATG 395

Qy 128 LeuSerAlaIleSerThrgluThrglyLeuSerlleleuThrpProileThrlyrHisCyS 147

Db 396 CTGAGGCGCATGACGACCGAGGCTGCCCTGTCATCTGTGGCCATCTGGACACTGC 455

Qy 148 ArgArgProArgTYrleuSerSerValMetCySValleuThrpAlaLeuSerleu 167

Db 456 CGCCGCCGACATGACTGTGCATCGGTGATGTGTCTGTGCTGTGGCCCTGTCTGCTG 515

Qy 168 ArgSerlleleuGluThrMePheCySAspPheleuPheSerGlyAlaAspSerValTrp 187

Db 516 CGAGATTCCTGAGATGATGTTCTGTGACTTCTGTTAGTGTGCTAATTCGTTGG 575

Qy 188 CySgluThrSerAspPheleThrlleAlaTrpleuValPheleuCySValleuCyS 207

Db 576 TGTGA-ACGTCAATTTATTAATCAATCGGCGGTGAT-TTTTATGTGTGGT-CTCTGT 632

Qy 208 GlySerSerleuValleuValArglleuCySglySerArglyMeProleu 226

Db 633 GGGTCCAG-CTGTGCTGTGTGTGAGAT-CTCTGTGATC-CGAGATGACAGCTG 686

RESULT 5

CN843258/c 825 bp mRNA linear EST 02-JUN-2004

LOCUS AGENCOURT.15669738 NIH_MGC_145 Homo sapiens cDNA clone

DEFINITION IMAGE:30706610 5', mRNA sequence.

ACCESSION CN843258

VERSION CN843258.1 GI:47948913

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS NIH-MGC http://mgs.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10N07 Bethesda, MD 20892

Email: cgabbs-remail.nih.gov

Tissue Procurement: GPCR Consortium

cDNA Library Preparation: GPCR Consortium

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov

Plate: IRB14 row: e column: 03

High quality sequence start: 21

High quality sequence stop: 561.

Location/Qualifiers

1..825

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30706610"

/tissue_type="mixed"

/lab_host="DH10B"

/clone_lib="NIH_MGC_145"

/notes="Vector: pCDNA3.1; Site 1: varies by clone; Site 2: varies by clone; ORFs were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-EcoRV-XmiI/XhoI-3', 5'-EcoRV-XmiI/NotI-3', EcoRV (TA cloned, non-directional). For information about which gene each clones represents, please visit our anonymous ftp site at ftp://image.llnl.gov/image/feareyed_plates/IRB1.presv.dat a Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 1.85e-89 Length: 825

Score: 978.00 Matches: 203

Percent Similarity: 86.33% Conservative: 18
 Best Local Similarity: 79.30% Mismatches: 29
 Query Match: 55.47% Indels: 7
 Gaps: 3

US-09-867-570-2 (1-337) x CN843258 (1-825)

```

QY      85 AAAAAAAAAA...leuPheleuSerGlyHisIleCysSerPro---leuArgLeuI 103
DB      794 GAGACAACTCTCTCTCTCCACAGGCGCCGCTTATTATTATCCCGGTAAAGTTCCA 735
QY      103 eaaNIleArg---HisProIleSerIys---IleLeuSerProValMetThrPheProIy 121
DB      734 TNCAGTATCCCGCCCACTCCATCTTAATAATCTTCTATC-CCTGTGATGATGTTTCTAC 676
QY      121 rPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuT 141
DB      675 CTTGGAGCCTGAGCTGTTTGTAGTGCGGTGAGCAGCAGCGCTGCTGCTGCTG 616
QY      141 pProIleTTrpTyrHisCysArg-ArgProArgTyrLeuSerSerValMetCysValLeu 161
DB      615 GCCCATCTGGTACCGCTGCCACACGCCCCACACACCTGTCAAGCGGTGTGTCTGTC 556
QY      161 eutTPAlaLeuSerLeuLeuArgSerIleLeuGluTTrpMetPheCysAspPheLeuPhe 181
DB      555 TCTGGAGCCTGTCCTCTCTGCGAGCAGTCTCGAATGATGATGATGAGCTTCTGTTCA 496
QY      181 erGlyAlaAspSerValTTrpCysGluThrSerAspPheIleThrIleAlaTTrpLeuValP 201
DB      495 GTGGTGTGATTCCTGCTGGTGTGATCAAGTCAAGTTCATCAAGTCCGCTGCTGATTT 436
QY      201 heLeuCysValValLeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGly 221
DB      435 TTTTATGTGTGCTCTGTGTGCTGACCTGTGCTGCTGATCAGATTCCTGTGTGAT 376
QY      221 eArGlyMetProLeuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPhe 241
DB      375 CCGGAAAGATACCGCTACGCTAGCGTACGACATCTGTCTACAGTATGCTTCTCC 316
QY      241 euLeuCysGlyLeuProPheGlyIleGlnTTrpAlaLeuPheSerArgIleHisLeuAsp 261
DB      315 TCCCTGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 256
QY      261 rPlyValLeuPheCysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSer 281
DB      255 GGGAAAGCTTATTTGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 196
QY      281 eAlaAspProIleIleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgGln 301
DB      195 GTGCGAAGCCCATCATTTACTTCTTCTGCGCGCTCTTATGAGCAGCGTCAAAATAGG 136
QY      301 snLeuValLeuValLeuGlnArgAlaLeuGlnAspThrPro-GluValAspGluGly 320
DB      135 ACCGAAAGCTGCTTCTCAGAGGCTCTGAGAGAGCGGTCTGCAAGTGAATGAAGTGA 76
QY      321 GlyTTrpLeuProGlnGluThrLeuGluLeuSerGlySerArg 334
DB      75 GGGCAGCTCTCTGAGAAATCTGAGAGCTGTGCGGAGAGAGA 34
  
```

RESULT 6
 CN843633/c 835 bp mRNA linear EST 02-JUN-2004
 LOCUS AGENCOURT_15658848 NIH_MGC_145 Homo sapiens cDNA clone
 DEFINITION IMAGE:30706619 5', mRNA sequence.
 ACCESSION CN843633
 VERSION CN843633.1 GI:47949288
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 835)
 AUTHORS NIH-MGC <http://mhc.ncl.nih.gov/>.

TITLE JOURNAL
 COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabs-remail.nih.gov
 Tissue Procurement: GPCR Consortium
 CDNA Library Preparation: GPCR Consortium
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.llnl.gov>
 Plate: IRB14 row: e column: 12
 High quality sequence stop: 660.
 Location/Qualifiers
 1. 835
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30706619"
 /issue_type="mixed"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_145"
 /note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2:
 varies by clone; ORFs were PCR-amplified and cloned into
 pCDNA3.1 by the GPCR Consortium. Cloning sites vary by
 clone and include the following: 5'-EcoRV-XbaI/XhoI-3',
 5'-EcoRV-XbaI/NotI-3', EcoRV (TA cloned, non-directional).
 For information about which gene each clone represents,
 please visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRB14.presv.dat
 A Note: this is a NIH_MGC Library."

FEATURES
 source

ORIGIN

Alignment Scores:
 Pred. No.: 3,036-89 Length: 835
 Score: 976.00 Matches: 198
 Percent Similarity: 86.85% Conservative: 20
 Best Local Similarity: 78.88% Mismatches: 30
 Query Match: 55.36% Indels: 6
 Gaps: 1

US-09-867-570-2 (1-337) x CN843633 (1-835)

```

QY      88 PheLeuPheLeuSerGlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHis 107
DB      822 TTCCTTACAGCTTTCCAA-GATTATACGTTTCCGCAATTACGCTTTCATCAATATT-CAG 765
QY      108 ProIleSerIysIleLeuSerPro-----ValMetThrPheProTyrPheIleGly 124
DB      764 CCATCTCATCCCGCAAA-AATCCCTGCTGTTCTGTGATGACCTTCCCTTACAGGC 706
QY      125 LeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTTrpProIleTTrp 144
DB      705 CTAGTATGTGTGAGCGGCATCAGCAGCGGCTGCTGTCTGTCTGTGAGCCCATGTGG 646
QY      145 TyrHisCysArgArgProArgTyrLeuSerSerValMetCysValLeuLeuTTrpAlaLeu 164
DB      645 TACCGTCCCGCGCGCCACACACCTGTACAGCGCTGTGTCTCTGCTGCGGCGCTG 586
QY      165 SerLeuLeuArgSerIleLeuGluTTrpMetPheCysAspPheLeuPheSerGlyAlaAsp 184
DB      585 TCCCTGTGTTTGTATGATGAGTGAAGGTTCTGTGACATTCCTGTTAAGGTCGTAT 526
QY      185 SerValTTrpCysGluThrSerAspPheIleThrIleAlaTTrpLeuValPheLeuCysVal 204
DB      525 TCTAGTTGGTGTAAACGTCAAGATTTTCATCCAGTCCGCGGCTGATTTTATGTGTG 466
QY      205 ValLeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgTyrMet 224
DB      465 GTTCTGTGTGTTTCCAGCTGCTGCTGCTGTGTCAGATCTCTGTGATCCCGAAGATG 406
  
```

QY	225	ProleutrraAgLeuTyValThrlleuLeuTrValleValPheLeuLeuCy61y	244
Db	405	CCGCGACACAGGCTCTACTGACCACTCTGCTCAAGTCGCTGCTTCTCTCTCGCGGC	346
QY	245	LeuProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAspTrpIlyValIleu	264
Db	345	CTGCCCTTCGGCATTCCTGGGGGCCCTTAATTACAGAGTGCACCTTAATTGGAAAGTCTTA	286
QY	265	PheCysHisValHisLeuValSerIlePheLeuSerAlaLeuAspSerPheAlaAspPro	284
Db	285	TATTCATCATGTTATCTGGTTTCATGTCATGCCCTGCTCTCTTAACACAGTAGTCCAACCC	226
QY	285	IleIleTyPhePheValAlGlySerPheArgGlnArgGlnAsnArgGlnAsnLeuIlySeru	304
Db	225	ATCATTAATCTCTTCTGTGGGCTCCCTTTAGCGACGCTCAAAATAGGCAACACTGAAGCTG	166
QY	305	ValLeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyIleGlyTrpLeuPro	324
Db	165	GTTCTCCAGAGGGCTCTGACAGACAAAGCTGAGTGTGAATGAAGGAGGACGCTTCT	106
QY	325	GlnGluThrLeuGlnLeuSerGlySerArgLeu	335
Db	105	GAGGAAGCTGAGACTGTGGGAAAGACAGATTG	73
RESULT 7			
LOCUS	CN831885/c	827 bp	mRNA linear EST 02-JUN-2004
DEFINITION	AGNCNCORRT.15669850 NIH_MGC.145 Homo sapiens cDNA clone		
IMAGE:	30706617 5', mRNA sequence.		
ACCESSION	CN831885		
VERSION	CN831885.1	GI:47935638	
KEYWORDS	EST.		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens (human)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euteheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 827)		
JOURNAL	NIH-MGC http://mgs.nci.nih.gov/.		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NTH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cga@bs-remail.nih.gov Tissue Procurement: GPCR Consortium cDNA Library Preparation: GPCR Consortium cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov Plate: IRB14 row:e column: 10 High quality sequence scop: 653. Location/Qualifiers 1. 827 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:30706617" /tissue_type="mixed" /lab_host="DH10B" /note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2: varies by clone; ORFs were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-EcoRV-XbaI/XhoI-3', 5'-EcoRV-XbaI/NotI-3', EcoRV (TA cloned, non-directional). For information about which gene each clone represents, please visit our anonymous ftp site at ftp://image.lnl.gov/image/rearrayed_plates/IRB1.presv.dat a Note: this is a NIH-MGC Library."		

[illegible]

ORIGIN

/clone_11b="Atherys RAGE Library"
 /note="See 'Creation of Genome-wide Protein Expression
 Libraries using Random Activation of Gene Expression',
 Nature Biotechnology, in press. Note that even though the
 cell type indicated is Htt1080, since a random activation
 method was used, these sequence tags are not necessarily
 expressed in Htt1080 under normal circumstances."

Alignment Scores:

Pred. No.:	2.44e-81	Length:	787
Score:	898.50	Matches:	190
Percent Similarity:	84.34%	Conservative:	20
Best Local Similarity:	76.31%	Mismatches:	39
Query Match:	50.96%	Indels:	4
DB:	4	Gaps:	0

US-09-867-570-2 (1-337) x BG208126 (1-787)

```

Qy 81 LeuAnLeuValAlaAlaSpHeuPheLeuSerGlyHisIleIleCysSerProLeu 100
Db 747 CTGACGCTGGGGGGGGGACCTTCTCTTCCACAGCCACGATCCA-TCCCTGTTA 689
Qy 101 ArgLeuIleAsnIleArgHisProIleSerIleLeuSerProValMetThrPhePro 120
Db 688 CGCCATCAATAGTCCATCCCATCTCTTAATCTTTAGCTGTGATGACCTTTCC 629
Qy 121 TyrPheIleGlyLeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeu 140
Db 628 TACTTGCGACGGCTGATATCTGAGTGCATGAGCAGCAGGCGCTGCTCAATCTG 569
Qy 141 TrpProIleTrpTyrHisCysArgArgProArgTyrLeuSerSerValMetCysValLeu 160
Db 568 TGGCCATCTGTAAAGCTGCGGCCACCCCA--CACTCCGCAAGGTCGTGTGCTCG 511
Qy 161 LeuTrpAlaLeuSerIleLeuArgSerIleLeuGluTyrTrpMetCysAspPheLeuPhe 180
Db 510 CTCTGGGCGCTCTCCCTGCTGTAAGACATCTCGAGTGAATGTTCTGTGACTCCCTGTT 451
Qy 181 SerGlyAlaAspSerValTyrCysGluThrSerAspPheIleThrIleAlaTrpLeuVal 200
Db 450 AGTGAATGCTGATCTGTGTGGTGTCAACATCAGA-TTCAATCAAGTTACGTGCTGAT 392
Qy 201 PheLeuCysValIleLeuCysGlySerSerIleValIleLeuValArgIleLeuCysGly 220
Db 391 TTTTATTGTGTGCTCTGTGTGCGCCAGCTGTGCTTACTGATGATCCTCTGTGA 332
Qy 221 SerArgIleMetProLeuThrArgLeuTyrValThrIleLeuThrValLeuValPhe 240
Db 331 TCCGAGAGATGCCACTGACAGGCTGTACATGACCAT-CTGCTCAAGTGTGCTGCTTC 273
Qy 241 LeuLeuCysGlyLeuProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAsp 260
Db 272 CTCTCTGTGGGCTGCGCCATGGCCATTCAGTGGGCTGTCTTCCAGGATCCACATGGAC 213
Qy 261 TrpIleValIlePheCysHisIleValHisLeuValSerIlePheLeuSerAlaLeuAsnSer 280
Db 212 TGGGAAGCTTATATCTCATGTTCTGCTTCCATTTTCTTGGCGTCTCTTAACAGC 153
Qy 281 SerAlaAsnProIleIleTyrPhePheValGlySerPheArgGlnArgGlnAsnArgGln 300
Db 152 AGTGCACACCCATCATTTACTTCTCATGGGTTTGTTAGCAGCATCAAAATTTAGCAG 93
Qy 301 AsnLeuValLeuValLeuGlnArgAlaLeuGlnAspThrProGluValaLapGluGlyGly 320
Db 92 AACCTGAAGCTGTTCTCCAGAGGGATCTGCAGAGACACGCTGAGTGAAGGTGA 33
Qy 321 GlyTrpLeuProGlnIleThrLeuGlu 329
Db 32 TGGTGTATCTTCAGGAACCTTGAG 6

```

RESULT 10
 CN839125

LOCUS CN839125 871 bp mRNA linear EST 02-JUN-2004
 DEFINITION AGENCOURT 15864180 NIH MGC 145 Homo sapiens cDNA clone
 IMAGE:30706609 3', mRNA sequence.
 ACCESSION CN839125
 VERSION CN839125.1 GI:47944780
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 871)
 NIH-MGC http://mgs.nci.nih.gov/.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contract: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgaabbs-r@mail.nih.gov
 Tissue Procurement: GPCR Consortium
 cDNA Library Preparation: GPCR Consortium
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:
 http://image.llnl.gov
 Plate: IRB14 row: e column: 02
 High quality sequence stop: 701.
 Location/Qualifiers

FEATURES

source

1..871
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30706609"
 /issue_type="mixed"
 /lab_host="DH10B"
 /clone_11b="NIH MGC 145"
 /note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2:
 varies by clone; ORFs were PCR-amplified and cloned into
 pCDNA3.1 by the GPCR Consortium. Cloning sites vary by
 clone and include the following: 5'-EcoRV-XmiI/XhoI-3',
 5'-EcoRV-XmiI/NotI-3', EcoRV (TA cloned, non-directional).
 For information about which gene each clones represents,
 please visit our anonymous ftp site at
 ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat
 A Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.:	1.31e-76	Length:	871
Score:	853.00 <td>Matches:</td> <td>172 </td>	Matches:	172
Percent Similarity:	82.43% <td>Conservative:</td> <td>11 </td>	Conservative:	11
Best Local Similarity:	77.48% <td>Mismatches:</td> <td>37 </td>	Mismatches:	37
Query Match:	48.38% <td>Indels:</td> <td>2 </td>	Indels:	2
DB:	7 <td>Gaps:</td> <td>0 </td>	Gaps:	0

US-09-867-570-2 (1-337) x CN839125 (1-871)

```

Qy 6 SerTrpValIleArgLeuGlyPheLeuSerMetAspSerThrIleProValLeuGlyThr 25
Db 195 TCCAGGAGTACACAGACTAGAGGTTTCTGAGCAGATCCAAACATCTCAACCTTGACACA 254
Qy 26 GlnLeuThrProIleAsnGlyArgGluGluTyrProCysTyrIleGlnThrLeuSerPhe 45
Db 255 GAACGTACACCAATCAACGGAACCTGAGAGACTCTTGTCTACACAGACCTTGAGCCTC 314
Qy 46 ThrGlyLeuThrCysIleValSerLeuValAlaLeuThrGlyAsnAlaValValLeuTrp 65
Db 315 ACGGTGCTACGTCATCTTCCCTTGTGCGGCTGACAGGAAAGCAAGTGTGCTCGG 374
Qy 66 LeuLeuGlyCysArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAla 85
Db 375 CTCTGGGCTGCGCATGCGCAGAGAACGCTTCTCATCTCATCTCACTTCACTTGGCGCGCA 434

```

QY	86	AlAsppPheLeuPheLeuSerGlyHisIleIleIleCysSerProLeuArgLeuIle	105
Db	435	GCAGATTCCTCTTCTCTCAGCGGCCGCTTAATATTCCTCTGTAAGCTTCAATCAGATATC	494
QY	106	ArgHisProIleSerIleIleLeuSerProValMetThrPheProIlyrPheIleGlyLeu	125
Db	495	CCCCATACCATCTCTAAATATCTCTATCTGTAGATGATGTTTTCCTACTTGCAGGCCGCG	554
QY	126	SerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTrpProIleTrpIlyr	145
Db	555	AGCTTTCTGAATGCGGTAGACCCAGGCGCTCGCTCGCTGCGGCCATCTGGTAC	614
QY	146	HisCysArgArgProArgTyrLeuSerSerValMetCysValLeuLeuTrpAlaLeuSer	165
Db	615	CGCTGCCACCGCCACACACACTGTGACGGGTGTGTGTCTGTCTGGGCCCTGTCTC	674
QY	166	LeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeuPheSerGlyAlaAspSer	185
Db	675	CTGCTGCCGAGACATCTCGAATGAGTGTATGTGCTTCTCTGTTCAGTGTCTGTATTTCT	734
QY	186	ValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuValPheLeuCysValVal	205
Db	735	GCTTGGTGTCCAACTCAAGATTTCATCACAGTCGGCTGTATTTTATATGTGTGGGT	794
QY	206	LeuCysGlySerSerLeuValLeuLeuValArgIleLeuCysGlySerArgIlyMetCp	225
Db	795	CTCTGGGGGTCCAGCGCTGCTCTCTGTATTCAGATCTCTGGGGATTCCCGAATAAAAC	854
QY	225	ro 225	
Db	855	CC 856	

RESULT 11				
CD051096				
LOCUS	924 bp	mRNA	linear	EST 09-MAY-2003
DEFINITION	CD051096			
	AGENCOURT_1398113 NIH_MGC_173 Homo sapiens cDNA 5', mRNA sequence			

SOURCE ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 924)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D.

FEATURES
SOURCE

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/issue_type="embryonic trophoblasts, made from WA01 stem
cells"
/lab_host="DH10B Tona"
/clone_id="NH1 MGC 173"
/notes="Vector: pDONR01; Site_1: attP2; Site_2: attP1;
libr_priming - oligo dt method - full-length enriched;
libr_provider - Bradfield"

```

ORIGIN

Alignment Scores:	
Pred. No.:	3,46e-73
Score:	820.00
Percent Similarity:	83.11%
Best Local Similarity:	78.54%
Query Match:	46.51%
DB:	6
	Gaps:
	924
	Matches:
	172
	Mismatches:
	10
	Mismatches:
	31
	Indels:
	6
	Gaps:
	0

US-09-867-570-2 (1-337) X CD051096 (1-924

Oy 8 ValIleArgLeuGIyPheLeuSerMetAspSerThrIleProValLeuGIyThrCluLeu 27
 Db 245 GTACACGACTGAGGGTTTCTGAGCATGGAATCCAAACCATCTCAACCTTGAGACACAGACTG 304
 Oy 28 ThrProIleAsnGIyArgJugIuThrProCysTyrIyGlnThrLeuSerPheThrGIy 47
 Db 305 ACACCAATTCACAGGAACGTGAGGAACTCTTGTCTAACAGACAGACTTGAGCTCCAGGCTC 364
 Oy 48 LeuThrCysIleValSerLeuValAlaLeuThrGIyAsnAlaValValLeuThrLeuLeu 67
 Db 365 CTGACGTGCATCGTTTCCCTGTGGGGGTGACAGGAAGACGGGTGTGCTCGGCTCTCG 424
 Oy 68 GIyCysArgMetArgArgAsnAlaValSerIleTyrIleLeuAsnLeuValAlaIAsp 87
 Db 425 GGCTGCGCGCAAGCCGACGGAAGCGCTTCTCCATCTACACCTTCMACTTGCGCGACAGAC 484
 Oy 88 PheLeuPheLeuSerGIyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHis 107
 Db 485 TTCTCTTCTCCACACGGCGGCTATATATATCTCCGTGTAACCTTCATGAGATCCCCCAT 544
 Oy 108 ProIleSerIyIleLeuSerProValMetThrPheProTyrPheIleGIyLeuSerMet 127
 Db 545 ACCATCTCTAAATCTCTCTATCTGTGATGATGTTTCTTCTTCTTGACGGCTGAGCTTT 604
 Oy 128 LeuSerAlaIleSerThrGIyArgCysLeuSerIleLeuTyrProIleTyrIleCys 147
 Db 605 CTGAGTGCGCGAGACACGAGCGGTGCGTCCGTCGTCGAGCCCATCTGGACCGCTGC 664
 Oy 148 ArgArgProArgTyrLeuSerSer-ValMetCysValLeuLeuThrPalalaLeuSerLeu 167
 Db 665 CACGCGCCCAACCTCTGCACGCGAGATGTGTCTCTGCGGGCCCTGCTCTCT 724
 Oy 167 uArgSerIleLeuGIuThrMetPheCysAspPheLeuPheSerGIyAlaAspSerValTyr 187
 Db 725 GCGAGAGATCTCGAGTGGAGTATTATGTGCTTCTCTGTACGTGGTGTGATTCGCTTG 784
 Oy 187 pCysGIuThrSerAspPheIleThrIleAlaTyrLeuVal-PheLeuCysVal-ValLeu 206
 Db 785 GTGCAACAACATCATATTTCATCAGATCGCGCTGCTGATTTTTTTATGTGTGGGTTCTC 844
 Oy 207 CysGIySerSerLeu-ValLeu-LeuValArgIle-LeuCysGIy 220
 Db 845 TGTGGGTCCAGCCTGGGTCTCGCTGACAGATTTCTGTGTGG 889

RESULT 12	871 bp	mrna	linear	EST 21-APR-2001
BG210740/c				
LOCUS	BG210740			
DEFINITION	RS310287	Athersys RAGE	Library Homo sapiens	cDNA, mRNA sequence.
ACCESSION	BG210740			
VERSION	BG210740.1	GI:13732427		
KEYWORDS	EST,			

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 871)

Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R., Cain, S., Beventhal, C., Thornton, M., Ramachandran, R., Whittington, J., Letner, L., Costanzo, D., McGill, K., Booser, S., Mays, R., Smith, S., Veloso, N., Kikka, A., Hells, J., Cottrill, K., Lo, K., Offenbacher, J., Danzig, J., and Nucari, M.

TITLE

Creation of genome-wide protein expression libraries using random

JOURNAL
 MEDLINE
 PUBMED
 COMMENT

activation of gene expression
 Nat. Biotechnol. 19 (5), 440-445 (2001)
 21227151
 11329013
 Contact: Scott J. Cain
 Atherys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scain@atherys.com
 High quality sequence stop: 456.
 Location/Qualifiers
 1..871
 /organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 /cell_line="HT1080"
 /clone_id="Atherys RAGE Library"
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN

Alignment Scores:
 Pred. No.: 1.03e-72 Length: 871
 Score: 815.00 Matches: 180
 Percent Similarity: 68.14% Conservative: 21
 Best Local Similarity: 61.02% Mismatches: 66
 Query Match: 46.23% Indels: 29
 DB: 4 Gaps: 4

US-09-867-570-2 (1-337) x BG210740 (1-871)

Qy 54 LeuValAlaLeuThrGlyAsnAlaValLeuTrpLeuGlyCysArgMetCArg 73
 Db 867 ATCATAGCTTACTGTGTACACCTACCCGCGTGGCGGCGGCGTCCGCCG 808
 Qy 74 AsnAlaValSerIle-----78
 Db 807 ACAGCCTGCGGCGTAAAGTATTCCTCTGTAGACCCCTCATGGGCTGTCGAT 748
 Qy 79 -----TyrIleLeuAsnLeuValAlaAlaAspPheLeu 91
 Db 747 GGGGACCAAGCCGCTCTGTATAGACCGCTATATCCGAGAAATCTTTTCCC 668
 Qy 92 SerGlyHisIleIleCysSerProLeuArgLeuIleAsnIleArgHisProIleSer 111
 Db 687 AGGCCCCGTATATGCGGTAGACTTCG-----AGTTCCCAAACTCTCT 640
 Qy 112 IleLeuSerProValMetThrPheProTyrPheIleGlyLeuSerMetLeuSerAlaIle 131
 Db 639 AAAACCTATCTGTCAGATGTTTCT---ACTTCAGGCGTGAAGCTTTTAAATGCGGTAG 583
 Qy 132 SerThrGlyArgCysLeuSerIleLeuTrpProTyrHisCysArgArgProArg 151
 Db 582 CACGAAACCCGC---CTGTCCGCTCTGTG---CCCATCTGGTCCGTAGCTACCGCCACCA 527
 Qy 152 TyrLeuSerSerValMetCysValLeuLeuTrpAlaLeuSerLeuLeuArgSerIleLeu 171
 Db 526 CACCTGTACCGGTGTGTGTCTGTCTGTGGCCCTGTCCNTGTGCGGAGACATCTG 467
 Qy 172 GluTrpMetPheCysAspPheLeuPheSerGlyAlaAspSerValTrpCysGluTrpSer 191
 Db 466 GAGTGTATGTATGTGCTTCCCTGTTCAAGGTGTGATTTCTGTTGTTCAAACTCA 407
 Qy 192 AspPheIleThrIleAlaTrpLeuValPheLeuCysValValLeuCysGlySerSerLeu 211
 Db 406 GATTTTCATACAGCCGCTGCTGATTTTATGTTGTTCTCTGTGGGTCCAGCTG 347
 Qy 212 ValLeuLeuValArgIleLeuCysGlySerArgIleMetProLeuThrArgLeuVal 231

Db 346 GTCTGCTGATCAGGATTTCTGTGGATCCGGAGATACCGCTACACGAGCTGACGTG 287
 Qy 232 ThrIleLeuLeuThrValLeuValPheLeuLeuCysGlyLeuProPheGlyTlIleTrp 251
 Db 286 ACCATCTGCTCAGACTGCTGCTTCTCTCTGTGGCTGCTCCCTTTGGCATTCAGTTT 227
 Qy 252 AlaLeuPheSerArgIleIleLeuAspTrpIleValLeuPheCysHisValHisLeuVal 271
 Db 226 TTCTTATTTTATGATTCACGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 167
 Qy 272 SerIlePheLeuSerAlaLeuAsnSerSerAlaAsnProIleIleTyrPhePheValGly 291
 Db 166 TCATTTCCCTGCTCCGCTCTTAAACAGAGTCGACCAACCATCATTTACTTCTGCGTGGC 107
 Qy 292 SerPheArgGlnArgGlnAsnArgGlnAsnArgGlnAsnArgGlnAsnArgGlnAsn 311
 Db 106 TCTTTAGGCAAGCTCAAAATAGGAGAACTGAAGCTGTTCTTCAGAGGAGGCTTGCAG 47
 Qy 312 AspThrProGluValAspGluGlyGlyTyrTrpLeuProGlnGlu 326
 Db 46 GACGCTGAGAGTGAATGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2

RESULT 13
 CN839765/c 953 bp mRNA linear EST 02-JUN-2004
 LOCUS
 AGENCOURT 15864292 NIH MGC 145 Homo sapiens cDNA clone
 DEFINITION
 IMAGE:30706616 3', mRNA sequence.
 ACCESSION
 CN839765
 VERSION
 CN839765.1 GI:47945420
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 953)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: c9abbs-remail.nih.gov
 Tissue Procurement: GPCR Consortium
 cDNA Library Preparation: GPCR Consortium
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ULNL at:
 http://image.llnl.gov
 plate: IRB4 row: e column: 09
 High quality sequence stop: 578.
 Location/Qualifiers
 1..953

FEATURES

source

/organism="Homo sapiens"
 /mol_type="RNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30706616"
 /issue_type="mixed"
 /lab_host="DH10B"
 /note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2: varies by clone; ORFs were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-EcoRV-XmiI/XhoI-3', 5'-EcoRV-XmiI/NotI-3', EcoRV (TA cloned, non-directional). For information about which gene each clones represents, please visit our anonymous ftp site at ftp://image.llnl.gov/image/fearrayed_plates/IRB1.presv.dat a Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pred. No.: 3,02e-68 Length: 953
 Score: 772.00 Matches: 156
 Percent Similarity: 87.24% Conservativeness: 15
 Best Local Similarity: 79.59% Mismatches: 23
 Query Match: 43.79% Indels: 2
 DB: 7 Gaps: 0

US-09-867-570-2 (1-337) x CN839765 (1-953)

QY 142 ProleuTyrHisCysArgArgProArgTyrLeuSer-SerValMetCysValleu 161
 Db 675 CCAATTGGTACCCGCGTCCGCCCCACACAACCAAGCTACGCGCTGTGTCTCTAC 616
 QY 161 euTTPAlaLeuSerLeuLeuArgSerileuGluTTPMetPheCysAspPheLeuPhe 181
 Db 615 TCTGGGCGCTGTCTCTGTCTGTATAGTGTGAGTGTGAGGTTCTGTGACTTCTGTATTA 556
 QY 181 exGlyAlaAspSerValTTPCysGluThrSerAspPheIleThrIleAlaThrLeuValP 201
 Db 555 GTGGTGTGATTCAGTTGGTGTGAACGTCAGATTTCCTCCAGTCGCGGTGATTT 496
 QY 201 heluCyValAlaLeuCyGlySerSerLeuValleuValArgIleuCyGlyS 221
 Db 495 TTTTATGTGTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 436
 QY 221 exArgLysMetProleuThrArgLeuTyrValIleLeuLeuThrValleuValPhe 241
 Db 435 CCGGAGAGATGCCGCTGACGAGCTGTACGATCCATCTCTCTCAAGTGTGTCTTCC 376
 QY 241 euLeuCyGlyLeuProPheGlyIleGluTTPAlaLeuPheSerArgIleIleLeuApt 261
 Db 375 TCCCTGTGGGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 316
 QY 261 rplyValLeuPheCysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSers 281
 Db 315 TGGAGGCTTATATGTCACTTATCTGTTCGATCTGTCTGTCTGTCTGTCTGTCTGT 256
 QY 281 exAlaAspProIleIleTyrPhePheValGlySerPheArgIleArgIleAsnArgIle 301
 Db 255 GTGGCAACCCCATCATTTACTTCTGTGGGCTCTTTAGGACAGTCGCAAAATAGGAGA 196
 QY 301 snLeuValLeuValLeuGlnArgAlaLeuGlnAspThrProGluValAspGluGly 321
 Db 195 ACCGAAAGCTGTCTCTCAAGAGGCTCTGACAGCAAGCCCTGAGTGTGATTAAGTGAAG 136
 QY 321 lYrTleuProGluGluThrLeuGluLeuSerGlySerArgLeu 335
 Db 135 GGCAGCTTCTGAGGAAAGCTGAGAGCTGTGGAGAGCAATTTG 92

RESULT 14
 CN837763 1009 bp mRNA linear EST 02-JUN-2004
 LOCUS ABEENCOURT 15864212 NIH MGC.145 Homo sapiens cDNA clone

DEFINITION IMAGE:30706611 3', mRNA sequence.
 ACCESSION CN837763

VERSION CN837763.1 GI:47943418
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1009)
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabs-remail.nih.gov
 Tissue Procurement: GPCR Consortium
 cDNA Library Preparation: GPCR Consortium

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: IRB14 row: e column: 04
 High quality sequence stop: 464.
 Location/Qualifiers
 1. 1009
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30706611"
 /tissue_type="mixed"
 /lab_host="DH10B"
 /clone_id="NIH MGC 145"
 /note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2:
 varies by clone; ORFs were PCR-amplified and cloned into
 pCDNA3.1 by the GPCR Consortium. Cloning sites vary by
 clone and include the following: 5'-BCORV-XmnI/XhoI-3',
 5'-BCORV-XmnI/NotI-3', BCORV (TA cloned, non-directional).
 For information about which gene each clones represents,
 please visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat
 A Note: this is a NIH MGC Library."

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 5.28e-68 Length: 1009
 Score: 770.00 Matches: 148
 Percent Similarity: 92.49% Conservativeness: 12
 Best Local Similarity: 85.55% Mismatches: 13
 Query Match: 43.68% Indels: 0
 DB: 7 Gaps: 0

US-09-867-570-2 (1-337) x CN837763 (1-1009)

QY 165 SerLeuLeuArgSerIleLeuGluTTPMetPheCysAspPheLeuPheSerGlyAlaAsp 184
 Db 560 TCCGAGGTGGGAGCAATCAAGATTTCATCACTACGAGTGTGATTTTATATGTGTG 501
 QY 185 SerValTTPCysGluThrSerAspPheIleThrIleAlaThrLeuValPheLeuCyVal 204
 Db 500 TTTCCTGTGTCAAAATCAAGATTTCATCACTACGAGTGTGATTTTATATGTGTG 441
 QY 205 ValLeuCyGlySerSerLeuValleuValArgIleLeuCyGlySerArgLysMet 224
 Db 440 GTTCTGTGTGGTCAACCTGTCTGTCTGTATCAAGATTCTGTGTATCCCGGAAGATA 381
 QY 225 ProleuThrArgLeuTyrValIleLeuLeuThrValleuValPheLeuCyGly 244
 Db 380 CCGGTGACTAGGCTGTACGATCCATCTCTCACTACAGTGTGTCTCTCTGTGTGGC 321
 QY 245 LeuProPheGlyIleGluTTPAlaLeuPheSerArgIleHisLeuAspTTPlyValleu 264
 Db 320 CTGCGCTTTGGCATTCAGTTTCTCTATTTTATGATTCACAGTGAAGGAGTCTTA 261
 QY 265 PheCysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSersSerAlaAspPro 284
 Db 260 TTTTGTATGTTCATCTAGATTTCATTTTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 201
 QY 285 IleIleTyrPhePheValGlySerPheArgIleArgIleAsnArgIleAsnLeuLysLeu 304
 Db 200 ATCATTTACTTCTCTGTGGGCTCTTTAGGACAGGTCAAAATAGGACGAACCTGAACCTG 141
 QY 305 ValleuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGlyTTPLeuPro 324
 Db 140 GTTCTCAGAGGCTGTGACAGACGCTGTGAGTGTGATTAAGGAGGAGGAGCTTCT 81
 QY 325 GlnGluThrLeuGluLeuSerGlySerArgLeuGluGln 337
 Db 80 GAGGAATCTCTGAGAGCTGTGGGAGAGCAATTTGATCAG 42


```
QY 124 G|YLeuSerMetLeuSerAlaIleSerThrGluArgCysLeuSerIleLeuTrpProIle 143
Db 434 GGTCTGAGCAATCTCAGACACCATTAAGCATTAAGCGCTGCTGTGAATATAGCCATT 493
QY 144 TrpTyrHisCysArgProArgPheLeuSerSerValMetCysValLeuLeuTrpAla 163
Db 494 TGGTATGCGCTGAACCTCCAGACACACATCACTATCAATGTTTGTGCTTTGGGTT 553
QY 164 LeuSerLeuLeuArgSerIleLeuGluTrpMetPheCysAspPheLeuPheSerGlyAla 183
Db 554 ATGTCTTATTTGGGTCTCTCGAAGGAGGAGCATGTGGCTTACTGTTTATATAGCTTT 613
QY 184 AspSerValTrpCysGluThrSerAspPheIleThrIleAlaTrpLeuValPheLeuCys 203
Db 614 GACTCTTATTTGGTGTGAACATTGATGTTATCAATATATATGTCAGTTGTTTTTTT 673
QY 204 ValValLeuCysGlySerSerLeuValIleuValArgIleLeuCysGlySerArgLys 223
Db 674 GGTCTTCTGTGGGCTTAAGCTTCACTGCTGTCTGTGAGATCTTCTGTGGCTCACAGCA 733
QY 224 MetProLeuThrArgLeuTyrValThrIleLeuLeuThrValLeuValPheLeuLeuCys 243
Db 734 ATTCCTATGACACAGCTGTATGTACTATTATCACTCAAGCTTGTGCTTCTGATCTTT 793
QY 244 GlyLeuProPheGlyIleGlnTrpAlaLeuPheSerArgIleHisLeuAspTrpLysVal 263
Db 794 GGTCTTCCCTTTGGGATCTATGTGATCACTATCACTGATTAACAAATTTTATTATGTT 853
QY 264 LeuPheCysHisValHisLeuValSerIlePheLeuSerAlaLeuAsnSerSerAlaAsn 283
Db 854 GAATTTGTATTTTATCTTGAATATCTATCTATCTGTGTTAACAGCTGTATGAC 913
QY 284 ProIleIleTyrPhePheValGlySerPheArgGlnArgLys--AsnArgGlnAsnLeu 302
Db 914 CCATCATTTATTTCTTGTGGCTTCCATTAAGCACCGAAGTTCAGGCGGAAGACTCTC 973
QY 303 LysLeuValLeuGlnArgAlaLeuGlnAspThrProGluValAspGluGlyGly 320
Db 974 AAGCTACTTCTGACAGAGCCATGCAAGACACCCCTGAGAGAAACAAGTGA 1027
```

Search completed: December 2, 2004, 01:15:17
Job time : 2986 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 2, 2004, 00:13:59 ; Search time 34384 Seconds
(without alignments)
11858.184 Million cell updates/sec

Title: US-09-867-570-3

Perfect score: 8622
Sequence: 1 tghatgaagcccaatgctact.....gltccccaagggcccttacc 8622

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

GenEmb1:*
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_srs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8622	100.0	8622	6	AX429467 Sequence
2	8574.6	99.5	91923	9	AC108007 Homo sapi
3	8564.2	99.3	172939	9	AC090099 Homo sapi
4	5966.2	69.2	163718	9	AC023078 Homo sapi
5	5380.4	62.4	165490	2	AC139482 Homo sapi
6	4719	54.7	192618	2	AC020568 Homo sapi
7	4424.6	52.1	192618	2	AC020568 Homo sapi
8	4424.6	51.3	165490	2	AC139482 Homo sapi
9	2601.8	30.2	189230	9	AC103974 Homo sapi
10	1535.8	17.8	156839	9	AC107948 Homo sapi
11	1526.8	17.7	104547	9	AC079120 Homo sapi
12	1521.2	17.6	179216	2	AC074016 Homo sapi
13	1521.2	17.6	181150	2	AC007256 Homo sapi
14	1517.2	17.6	125001	9	AC126613 Homo sapi
15	1507.4	17.5	121520	9	AC006477 Homo sapi
16	1503.4	17.4	155021	9	HS774G10 Human DNA
17	1498.4	17.4	137091	9	AC135452 Homo sapi
18	1497.2	17.4	120591	9	AC116098 Homo sapi
19	1482.6	17.2	102918	9	AC079809 Homo sapi

C	20	1478.2	17.1	205954	9	AC008546 Homo sapi
	21	1470	17.0	145250	9	CNS01D78 Homo sapi
	22	1462	17.0	74649	9	BX293995 Human DNA
	23	1462	17.0	175582	2	AC016977 Homo sapi
	24	1462	17.0	331221	2	AL390881 Homo sapi
	25	1457	16.9	70466	9	AC068299 Homo sapi
	26	1452.8	16.8	193826	2	AL513479 Human DNA
	27	1450	16.8	134344	2	HS4773124 Homo sapi
	28	1448.4	16.8	211654	2	AL591103 Homo sapi
	29	1447.8	16.8	172543	9	AC106714 Homo sapi
	30	1447.2	16.8	113428	9	CR626893 Human DNA
	31	1447.2	16.8	122867	9	CR626925 Human DNA
	32	1447.2	16.8	119858	2	AC020688 Homo sapi
	33	1446.8	16.8	110000	2	AL583837_1 Homo sapi
	34	1446.8	16.8	112948	9	AL136092 Homo sapi
	35	1446.6	16.8	135245	9	AF172277 Homo sapi
	36	1445.6	16.8	227054	9	AF172277 Homo sapi
	37	1445.6	16.8	167358	2	AC010673 Homo sapi
	38	1445.6	16.8	178628	2	AC015715 Homo sapi
	39	1445.6	16.8	181511	9	AC087738 Homo sapi
	40	1443	16.7	112968	9	AL360172 Homo sapi
	41	1442.4	16.7	173086	2	AC093833 Homo sapi
	42	1441.2	16.7	173086	2	AC138065 Homo sapi
	43	1439.6	16.7	173086	2	AL591705 Homo sapi
	44	1438.2	16.7	90526	9	AC007567 Homo sapi
	45	1438.2	16.7	133052	9	AC092393 Homo sapi
						AL136325 Human DNA

ALIGNMENTS

RESULT 1	AX429467	8622 bp	DNA	linear	PAT 21-JUN-2002
LOCUS	AX429467	Sequence 3 from Patent WO0234914.			
DEFINITION	AX429467				
ACCESSION	AX429467.1	GI:21540764			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
	Homo sapiens (human)				
REFERENCE					
AUTHORS	Zhao, Q., Beasley, E.M. and Wei, M.H.				
TITLES	Isolated human g-protein coupled receptors, nucleic acid molecules encoding human gpcr proteins, and uses thereof				
JOURNAL	Patent: WO 0234914-A 3 02-MAY-2002;				
	PR CORP NY (US)				
FEATURES					
Source	Location/Qualifiers				
	1..8622				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Query Match	100.0%; Score 8622; DB 6; Length 8622;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 8622; Conservative	0; Mismatches	0; Indels	0; Gaps	0;	
Qy	1	TGTTAGAGCCCAATGCTCTTATACCAAAACAGGAAGGATATCAAAAAGAAAC	60		
	1	TGTTAGAGCCCAATGCTCTTATACCAAAACAGGAAGGATATCAAAAAGAAAC	60		
Db	1	TGTTAGAGCCCAATGCTCTTATACCAAAACAGGAAGGATATCAAAAAGAAAC	60		
Qy	61	TATAGACAGTATGATGATATATATATATATATATATATATATATATATATATAT	120		
	61	TATAGACAGTATGATGATATATATATATATATATATATATATATATATATATAT	120		
Db	61	TATAGACAGTATGATGATATATATATATATATATATATATATATATATATATAT	120		
Qy	121	CGAATCCAGCAGAT	180		
	121	CGAATCCAGCAGAT	180		
Db	121	CGAATCCAGCAGAT	180		
Qy	181	GTGAGATAGTAT	240		
	181	GTGAGATAGTAT	240		
Db	181	GTGAGATAGTAT	240		

Db 181 GTGAGATAGGTTAATACATCAAACTAATAATGATATACATCATTAACAGATTA 240
Qy 241 AAAAAAATATACATGATCATCTCAATGATGCTGAAAAAGATTGGAAAAATCTAAC 300
Db 241 AAAAAAATATACATGATCATCTCAATGATGCTGAAAAAGATTGGAAAAATCTAAC 300
Qy 301 ATTTCTTATGATTTAAACCTTCAGCAAAATGACATAGAAAAGACATCTTAATGTA 360
Db 301 ATTTCTTATGATTTAAACCTTCAGCAAAATGACATAGAAAAGACATCTTAATGTA 360
Qy 361 TAAAAAGCATATATGACGAGCCACAGCAAACTTAATGTAATGGGAAAAAGTTGAAA 420
Db 361 TAAAAAGCATATATGACGAGCCACAGCAAACTTAATGTAATGGGAAAAAGTTGAAA 420
Qy 421 CATGTCCTGGAATCTGGAACAAGACAGATGCTACTTTCACCTTCTATTCAACT 480
Db 421 CATGTCCTGGAATCTGGAACAAGACAGATGCTACTTTCACCTTCTATTCAACT 480
Qy 481 AGTAGTGAAGTTTATAGCCAGACATCAGACAAAGAAAAGAAATCAAGGACCCAAAT 540
Db 481 AGTAGTGAAGTTTATAGCCAGACATCAGACAAAGAAAAGAAATCAAGGACCCAAAT 540
Qy 541 CAATTAAGAGAGATGCAAACTGTCTCTGTTCACTGATGATGATTTGATATCTAGAAA 600
Db 541 CAATTAAGAGAGATGCAAACTGTCTCTGTTCACTGATGATGATTTGATATCTAGAAA 600
Qy 601 CCTTAAGCTCATCCAGAAAGCTCTAGAACTGATACATTAATTTGATTAAGTTTCAG 660
Db 601 CCTTAAGCTCATCCAGAAAGCTCTAGAACTGATACATTAATTTGATTAAGTTTCAG 660
Qy 661 ATACAAATCAAAATGATACAAAAATCAGTACATCTGATACCAAGACAGTACCAAGCTG 720
Db 661 ATACAAATCAAAATGATACAAAAATCAGTACATCTGATACCAAGACAGTACCAAGCTG 720
Qy 721 GAATCAATCAAGAACTCAAACTTTTCAATAGCTGTGTAATAATACTTAAGAAATTT 780
Db 721 GAATCAATCAAGAACTCAAACTTTTCAATAGCTGTGTAATAATACTTAAGAAATTT 780
Qy 781 CTTTACCAGAGGTGGAAGACCTCTACAAAGAAACTACAAACAGCTGACATCAATA 840
Db 781 CTTTACCAGAGGTGGAAGACCTCTACAAAGAAACTACAAACAGCTGACATCAATA 840
Qy 841 GATGACAAACAGAGTGAACATCCCATGCTCATGATGAGGTGGAATCAATATTGTTG 900
Db 841 GATGACAAACAGAGTGAACATCCCATGCTCATGATGAGGTGGAATCAATATTGTTG 900
Qy 901 AAAATGACCATATGTCGCAAAAGCAATCTCAAGTTCAATGCAATTCACCAAAATATCA 960
Db 901 AAAATGACCATATGTCGCAAAAGCAATCTCAAGTTCAATGCAATTCACCAAAATATCA 960
Qy 961 TCATCATCTTCAAGAACTAGAAAAAACTTAATAATTTGATTTGAAACAAACCA 1020
Db 961 TCATCATCTTCAAGAACTAGAAAAAACTTAATAATTTGATTTGAAACAAACCA 1020
Qy 1021 AAAAAAAT 1080
Db 1021 AAAAAAAT 1080
Qy 1081 GGCATCATCTTACCATCTTCAAACTATATCTACAGGCTTATATCAACCAATCATGAG 1140
Db 1081 GGCATCATCTTACCATCTTCAAACTATATCTACAGGCTTATATCAACCAATCATGAG 1140
Qy 1141 CACTGACATTAAGGACATAGGACATAGGACATAGGACATAGGACATAGGACATAGGAC 1200
Db 1141 CACTGACATTAAGGACATAGGACATAGGACATAGGACATAGGACATAGGACATAGGAC 1200
Qy 1201 CCAATTAATATATAGCACTGATTTTGAACAAGCAACCAAAACATTAAGTGGGAAAA 1260
Db 1201 CCAATTAATATATAGCACTGATTTTGAACAAGCAACCAAAACATTAAGTGGGAAAA 1260
Qy 1261 GACATTTATGTTAAACAAATGCTGAGATTATTTGCAAGCCACATGTTGAAAGAAATGAAA 1320
Db 1261 GACATTTATGTTAAACAAATGCTGAGATTATTTGCAAGCCACATGTTGAAAGAAATGAAA 1320

Qy 1321 CTGATCCCTTGCTCTCACTTAATACAAAAATGATACAAAGATGATCAAAAGCTTAAA 1380
Db 1321 CTGATCCCTTGCTCTCACTTAATACAAAAATGATACAAAGATGATCAAAAGCTTAAA 1380
Qy 1381 TCTGAGACTTAAACCATTAATAATTTCTAGAGATTAATCAGAAAAATGCTTCTAGACAT 1440
Db 1381 TCTGAGACTTAAACCATTAATAATTTCTAGAGATTAATCAGAAAAATGCTTCTAGACAT 1440
Qy 1441 TCACTTAGGCAAAAGCTTATGAGGACCAAGAACTTTCACAACTAAGCATCTGACTAAG 1500
Db 1441 TCACTTAGGCAAAAGCTTATGAGGACCAAGAACTTTCACAACTAAGCATCTGACTAAG 1500
Qy 1501 AAATGATAGGACTTAATTAATACTAATAAGCTTTGCGACAGCAAAACATCATTAAGAG 1560
Db 1501 AAATGATAGGACTTAATTAATACTAATAAGCTTTGCGACAGCAAAACATCATTAAGAG 1560
Qy 1561 AGCAAAACAGAACCAACCGAGTGAAGAAATCTTCAAACTAAGCATCTGACTAAG 1620
Db 1561 AGCAAAACAGAACCAACCGAGTGAAGAAATCTTCAAACTAAGCATCTGACTAAG 1620
Qy 1621 ACTAATATCCGGAATCCACAGAACTCAACAAATCAGCAAGAAAGAAACAAATCC 1680
Db 1621 ACTAATATCCGGAATCCACAGAACTCAACAAATCAGCAAGAAAGAAACAAATCC 1680
Qy 1681 CATGAAAGAGTGGGCTAAGGACATGAATAGCAATTCCTCAAAAGAAAGATTAACAATGGC 1740
Db 1681 CATGAAAGAGTGGGCTAAGGACATGAATAGCAATTCCTCAAAAGAAAGATTAACAATGGC 1740
Qy 1741 CAACAAACAGAAAAAATGCTTAATACATCACTAATGATTAAGGAAATGTAATCAACATG 1800
Db 1741 CAACAAACAGAAAAAATGCTTAATACATCACTAATGATTAAGGAAATGTAATCAACATG 1800
Qy 1801 TAATGGAATACCACTTACTCTGCAAGAAATGCTATTAATTTAAAAATCTAAAAATATA 1860
Db 1801 TAATGGAATACCACTTACTCTGCAAGAAATGCTATTAATTTAAAAATCTAAAAATATA 1860
Qy 1861 GATGTTGGTGGGCTGTGTGATTAAGGAACTTTTCACTGCTGTGGGAAATGTAATAC 1920
Db 1861 GATGTTGGTGGGCTGTGTGATTAAGGAACTTTTCACTGCTGTGGGAAATGTAATAC 1920
Qy 1921 TTGCGCAACCATATAGAAAAAGTGTGAAATTTCTTAAGAACTAATAAGTATGATGAC 1980
Db 1921 TTGCGCAACCATATAGAAAAAGTGTGAAATTTCTTAAGAACTAATAAGTATGATGAC 1980
Qy 1981 CATTTGATCAGCAATCCCATTAATATGATTAATTAATTAATTAATTAATTAATTAAT 2040
Db 1981 CATTTGATCAGCAATCCCATTAATTAATGATTAATTAATTAATTAATTAATTAATTAAT 2040
Qy 2041 ACAACTCAGCCATTAATAAAGATTAATAATGATGACATTCACAGCAATCTAAGATGAAATGG 2100
Db 2041 ACAACTCAGCCATTAATAAAGATTAATAATGATGACATTCACAGCAATCTAAGATGAAATGG 2100
Qy 2101 AGACCTTAATTTAAGTGGGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2160
Db 2101 AGACCTTAATTTAAGTGGGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2160
Qy 2161 TTAACAAGTGGGCTTAAGTGTGAGAGACAGAAAGCATTAAGTGAATTAAGTGAATCTGCG 2220
Db 2161 TTAACAAGTGGGCTTAAGTGTGAGAGACAGAAAGCATTAAGTGAATTAAGTGAATCTGCG 2220
Qy 2221 GGACTTGAAGGAGATGGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
Db 2221 GGACTTGAAGGAGATGGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
Qy 2281 GTACACTGCTCAGGTGATGAGTGAACCAAAATCTCAGAAATTAACCACTTAAGAACTTATC 2340
Db 2281 GTACACTGCTCAGGTGATGAGTGAACCAAAATCTCAGAAATTAACCACTTAAGAACTTATC 2340
Qy 2341 CATGGAAGCAACACCACTGTTCCCAAAATCCCAATGAAATTAATAATTAATTAATTAAT 2400
Db 2341 CATGGAAGCAACACCACTGTTCCCAAAATCCCAATGAAATTAATAATTAATTAATTAAT 2400

```
QY 2401 ATGATTTTATTCACAGAAATTTAAAGTTCACTGTTTCAGAGTTTAAATGAAGTA 2460
DB 2401 ATGATTTTATTCACAGAAATTTAAAGTTCACTGTTTCAGAGTTTAAATGAAGTA 2460
QY 2461 AGAATGAAAAGTGTAGCAAGTGTAGCTCTGACAAATGGAGCTCTGATATTTTCACTTG 2520
DB 2461 AGAATGAAAAGTGTAGCAAGTGTAGCTCTGACAAATGGAGCTCTGATATTTTCACTTG 2520
QY 2521 CATACACTTCTCTGGCAATTTGGAAAAGAAATGATACATGAAATATATCAACCATATGATA 2580
DB 2521 CATACACTTCTCTGGCAATTTGGAAAAGAAATGATACATGAAATATATCAACCATATGATA 2580
QY 2581 AAGAAACATCAAAAATTTGTGTAGGCAATTTGACGCTTGAATGTGCTCCATATATCTAC 2640
DB 2581 AAGAAACATCAAAAATTTGTGTAGGCAATTTGACGCTTGAATGTGCTCCATATATCTAC 2640
QY 2641 TTTTTCATTGGATATTAAGCCTCATATGATATGATTTGATTTGATGATGCTTAC 2700
DB 2641 TTTTTCATTGGATATTAAGCCTCATATGATATGATTTGATTTGATGATGCTTAC 2700
QY 2701 GCCCATATTTGATTTTAAATCAAGACAGCAATTTACTTACAGGAAGTTGAAACAAGATGG 2760
DB 2701 GCCCATATTTGATTTTAAATCAAGACAGCAATTTACTTACAGGAAGTTGAAACAAGATGG 2760
QY 2761 GACCTGATAGGAGGCTTAAATGATCTGATATGGAACAAGGCTCAAGATCATCTCAG 2820
DB 2761 GACCTGATAGGAGGCTTAAATGATCTGATATGGAACAAGGCTCAAGATCATCTCAG 2820
QY 2821 TTAGATTTTGTCTCAAAATACCTCTGGGCTCTGATTTGACCATATGCTCATACAGAA 2880
DB 2821 TTAGATTTTGTCTCAAAATACCTCTGGGCTCTGATTTGACCATATGCTCATACAGAA 2880
QY 2881 ATAAACAAGCTGTCAAGCATCTTGTGAAGCCTGTGATGCTCAACGCTTTGATTTCACT 2940
DB 2881 ATAAACAAGCTGTCAAGCATCTTGTGAAGCCTGTGATGCTCAACGCTTTGATTTCACT 2940
QY 2941 CCGTGAAGCATCTCTGATATTAAGCAACAAGAAAGTCTCTGAAGTCACTGAATCCCA 3000
DB 2941 CCGTGAAGCATCTCTGATATTAAGCAACAAGAAAGTCTCTGAAGTCACTGAATCCCA 3000
QY 3001 GAAAGGCTCTCTACCTTTAGCAACAAGGAGGCTTTCAACATGGAACAAGAAAGCAT 3060
DB 3001 GAAAGGCTCTCTACCTTTAGCAACAAGGAGGCTTTCAACATGGAACAAGAAAGCAT 3060
QY 3061 AAGGTAAGTACCAAGAACTCTCTTCTCAACAGTCAATTAATGATTTTGTCTGAAGATC 3120
DB 3061 AAGGTAAGTACCAAGAACTCTCTTCTCAACAGTCAATTAATGATTTTGTCTGAAGATC 3120
QY 3121 ATGTCTCTTATGCTTCAACCTTGTGTCTTCAACAGTCAATTAATGATTTTGTCTGAAGATC 3180
DB 3121 ATGTCTCTTATGCTTCAACCTTGTGTCTTCAACAGTCAATTAATGATTTTGTCTGAAGATC 3180
QY 3181 AGACAGAGACATGAGAGCTTCTTCAAGAACTGAGTGTCTCAACCCAAACCTGTGAG 3240
DB 3181 AGACAGAGACATGAGAGCTTCTTCAAGAACTGAGTGTCTCAACCCAAACCTGTGAG 3240
QY 3241 CTCTAAATGTGTCTCCCTCTTCTAATTTATCTCCCATATCACTCTCTTCAATCA 3300
DB 3241 CTCTAAATGTGTCTCCCTCTTCTAATTTATCTCCCATATCACTCTCTTCAATCA 3300
QY 3301 TTCAATGTGCTCTATGAGAGACTGCTGCTCTTACATTTATTTAACAGACAGGGAGC 3360
DB 3301 TTCAATGTGCTCTATGAGAGACTGCTGCTCTTACATTTATTTAACAGACAGGGAGC 3360
QY 3361 ATGAGGACATTTCTTCCAGAGTTGAAGTGTCTATGAGCCAGTTCTTTGTTTCACTTAC 3420
DB 3361 ATGAGGACATTTCTTCCAGAGTTGAAGTGTCTATGAGCCAGTTCTTTGTTTCACTTAC 3420
QY 3421 TTTTCAATTTATTTCTTCTTGTCTATCTGAAAAGTCTAAGAAATATGATGGCCCA 3480
DB 3421 TTTTCAATTTATTTCTTCTTGTCTATCTGAAAAGTCTAAGAAATATGATGGCCCA 3480
QY 3481 ATATTTAAGAGTCTTCAATGAGAAAGTATTTACAAAGATGACAGAGTTAAGGCTCAG 3540
DB 3481 ATATTTAAGAGTCTTCAATGAGAAAGTATTTACAAAGATGACAGAGTTAAGGCTCAG 3540
QY 3541 GATCTTAAGACAAATATATAGGAGCACTACTTCTCCCTTATGAGTGAACGAGCAAG 3600
DB 3541 GATCTTAAGACAAATATATAGGAGCACTACTTCTCCCTTATGAGTGAACGAGCAAG 3600
QY 3601 GAAAGACAGTTTACATTTGCTGCAATGCCATAGCTTATGCTTATAGGCTGTGAGAGCAT 3660
DB 3601 GAAAGACAGTTTACATTTGCTGCAATGCCATAGCTTATGCTTATAGGCTGTGAGAGCAT 3660
QY 3661 GAGAGGCAAGTGAAGAAAGCCTGCTGAGCAAGCAAGCAAGCAAGCTGATATGTT 3720
DB 3661 GAGAGGCAAGTGAAGAAAGCCTGCTGAGCAAGCAAGCAAGCAAGCTGATATGTT 3720
QY 3721 TGAATCTGTGTCTCCACAAATCTCATGTTGATTTGATTTTCCAAATGTTGAGAAAGG 3780
DB 3721 TGAATCTGTGTCTCCACAAATCTCATGTTGATTTGATTTTCCAAATGTTGAGAAAGG 3780
QY 3781 CTTGTGGAGATGATTTATGATCACGGGAGTGTGTTGCAATGATTTTAAACACCAT 3840
DB 3781 CTTGTGGAGATGATTTATGATCACGGGAGTGTGTTGCAATGATTTTAAACACCAT 3840
QY 3841 CCCCCTTGTGATTTGTTGTGTGTA.TACTGA.CGAGTCTCAATGAATCTAGTTGTTTAAA 3900
DB 3841 CCCCCTTGTGATTTGTTGTGTGTGTA.CGAGTCTCAATGAATCTAGTTGTTTAAA 3900
QY 3901 GCGTGTAGCACTCTCCCTCTCTCTTACTCTGCTCTCAACATGAGAGCCTGCTC 3960
DB 3901 GCGTGTAGCACTCTCCCTCTCTCTTACTCTGCTCTCAACATGAGAGCCTGCTC 3960
QY 3961 CCCCTTGTGCTTCAACAAGATGGAAGCTTCTGAGGCTCTCCCAAGAGCAAGAGCTGC 4020
DB 3961 CCCCTTGTGCTTCAACAAGATGGAAGCTTCTGAGGCTCTCCCAAGAGCAAGAGCTGC 4020
QY 4021 TATGCTCTTGTACAGTCTGAGAGCTATTAAGCAATTAACCAATTTCTTCAATTAAT 4080
DB 4021 TATGCTCTTGTACAGTCTGAGAGCTATTAAGCAATTAACCAATTTCTTCAATTAAT 4080
QY 4081 TCCAGTCTGAGATTTCTTTTATGCAATTTGAAATGAAATGAAATCAACAGACAGAG 4140
DB 4081 TCCAGTCTGAGATTTCTTTTATGCAATTTGAAATGAAATGAAATCAACAGACAGAG 4140
QY 4141 CCAGAGATGGAATCCCAAGTGTCTTCTGCTGTCTTCCAGTCTCTGCTGCTGCTC 4200
DB 4141 CCAGAGATGGAATCCCAAGTGTCTTCTGCTGTCTTCCAGTCTCTGCTGCTGCTC 4200
QY 4201 CCAAGTGTCTCAATTTCAACAAGAAACAGAAATTAAGAAATCCCACTGATGTGTACATA 4260
DB 4201 CCAAGTGTCTCAATTTCAACAAGAAACAGAAATTAAGAAATCCCACTGATGTGTACATA 4260
QY 4261 GAAAGCACTCTCTTGGAGTCAACAAGATTAAGAAATGAAATGAAATCCCTCATGG 4320
DB 4261 GAAAGCACTCTCTTGGAGTCAACAAGATTAAGAAATGAAATGAAATCCCTCATGG 4320
QY 4321 TAAATGAGATTCCTCTCAACCTTCTGTATCTCTTAAATCTGAGGCTTCTCATTC 4380
DB 4321 TAAATGAGATTCCTCTCAACCTTCTGTATCTCTTAAATCTGAGGCTTCTCATTC 4380
QY 4381 TGAATGATTCCTCTCAATTTCAAGTCTATCAAGCTTAAATGTTTGTGCTGTCTTC 4440
DB 4381 TGAATGATTCCTCTCAATTTCAAGTCTATCAAGCTTAAATGTTTGTGCTGTCTTC 4440
QY 4441 TCTACTGTCACTTTATGAGAAATGTTGATTTGTTTAAATGATAGAAATTAAT 4500
DB 4441 TCTACTGTCACTTTATGAGAAATGTTGATTTGTTTAAATGATAGAAATTAAT 4500
QY 4501 GTAATTTTAAAGAAATATGATTTTGTGTTGAAATTAAGTTGCTGATCTTAAATAA 4560
DB 4501 GTAATTTTAAAGAAATATGATTTTGTGTTGAAATTAAGTTGCTGATCTTAAATAA 4560
QY 4561 GACATGAAGAAATATCTTAAACAGAAAGTATATGTTGCTGCTGAGCTATAGTT 4620
DB 4561 GACATGAAGAAATATCTTAAACAGAAAGTATATGTTGCTGCTGAGCTATAGTT 4620
```

Db 4561 GACATGAGAGAAATATCTTAACAAGAAAGTAATGTTGCTCGGGTCACTAGGTT 4620
 Qy 4621 CTGAATCTACAGATTCAACAACTACAGAGAACTTTTCCAAAATTAAGGTGGG 4680
 Db 4621 CTGAATCTACAGATTCAACAACTACAGAGAACTTTTCCAAAATTAAGGTGGG 4680
 Qy 4681 GAGTGTGTATGATCTGAGACAGGTACAAACTGTATTTCTTTGTCTGATTAATTTCTGAAAA 4740
 Db 4681 GAGTGTGTATGATCTGAGACAGGTACAAACTGTATTTCTTTGTCTGATTAATTTCTGAAAA 4740
 Qy 4741 CTACATATTAACAAGAACTTATATAGCAATTTGTGATTTTCTGATTAATTTCTGAAAACTTT 4800
 Db 4741 CTACATATTAACAAGAACTTATATAGCAATTTGTGATTTTCTGATTAATTTCTGAAAACTTT 4800
 Qy 4801 AAATGATTTAATGTATCTGGAGAAAGTGATAGATTAATCAAAATCAATTAATTAAG 4860
 Db 4801 AAATGATTTAATGTATCTGGAGAAAGTGATAGATTAATCAAAATCAATTAATTAAG 4860
 Qy 4861 AAATGAGCATCTGAGATTTTGTGCTGCTGGGGTCTGAAAGAAATCCCTGTAAT 4920
 Db 4861 AAATGAGCATCTGAGATTTTGTGCTGCTGGGGTCTGAAAGAAATCCCTGTAAT 4920
 Qy 4921 ACACAAAAATGACATCTTGAGATCTGAATGAAGCTCCAAAGCATCATACAGAA 4980
 Db 4921 ACACAAAAATGACATCTTGAGATCTGAATGAAGCTCCAAAGCATCATACAGAA 4980
 Qy 4981 TTCCAAAAATTTGCTGCTCCCAATTTCTGAGAGTTGCCCTCATCTTTGTGATCTTACAT 5040
 Db 4981 TTCCAAAAATTTGCTGCTCCCAATTTCTGAGAGTTGCCCTCATCTTTGTGATCTTACAT 5040
 Qy 5041 GGTTCCAGCAGCATTAAGCATTCAGCTTATGAAAAAGAGAGAGAGAGAGGCT 5100
 Db 5041 GGTTCCAGCAGCATTAAGCATTCAGCTTATGAAAAAGAGAGAGAGAGAGGCT 5100
 Qy 5101 TTGCTCTTCTTAATTAATCCATGAGCAGAGACTTGTCTGTCACTTTGTGATCTTCC 5160
 Db 5101 TTGCTCTTCTTAATTAATCCATGAGCAGAGACTTGTCTGTCACTTTGTGATCTTCC 5160
 Qy 5161 ACTTAACAGCACTGCTCATGAGATCTCATCCAGCATCAAGAAAACTGGAGATGTGGTCT 5220
 Db 5161 ACTTAACAGCACTGCTCATGAGATCTCATCCAGCATCAAGAAAACTGGAGATGTGGTCT 5220
 Qy 5221 CTGTCGCTCTGATCATCTCAGAAAGTTATGTGACCAAAAAAGAAATCTTTGGGCA 5280
 Db 5221 CTGTCGCTCTGATCATCTCAGAAAGTTATGTGACCAAAAAAGAAATCTTTGGGCA 5280
 Qy 5281 ACCAGCAGTCTCTTCAAGCCCTGACCTGTCTGTATTTCTGTCACTCAATCAAGATTTTCA 5340
 Db 5281 ACCAGCAGTCTCTTCAAGCCCTGACCTGTCTGTATTTCTGTCACTCAATCAAGATTTTCA 5340
 Qy 5341 GGAATCTCTCAGAAATTAATTAATGTGTGGGAGAGAACTGAGTCTGTCAGGA 5400
 Db 5341 GGAATCTCTCAGAAATTAATTAATGTGTGGGAGAGAACTGAGTCTGTCAGGA 5400
 Qy 5401 CTCAGAGGACAGGGGCTGGTATTTGACCTGCTCTTCACTGTTGAAACAGGAAACCT 5460
 Db 5401 CTCAGAGGACAGGGGCTGGTATTTGACCTGCTCTTCACTGTTGAAACAGGAAACCT 5460
 Qy 5461 TTAATTTCTAGGCTTACCTTATCTTATGTATATGAGATTAATCAATAGACGTCT 5520
 Db 5461 TTAATTTCTAGGCTTACCTTATCTTATGTATATGAGATTAATCAATAGACGTCT 5520
 Qy 5521 TTAAGAAATCATATAGATTTTAAACAATGCTTAATTTGTGTGATTAACAAGTAAAA 5580
 Db 5521 TTAAGAAATCATATAGATTTTAAACAATGCTTAATTTGTGTGATTAACAAGTAAAA 5580
 Qy 5581 AGACAGGACATGATCTTACCTTACGATCTTGGGTTTCAAGAGAAACAATATCA 5640
 Db 5581 AGACAGGACATGATCTTACCTTACGATCTTGGGTTTCAAGAGAAACAATATCA 5640
 Qy 5641 TACCATTAATTAATGATGACAAACAGTTTATGTCTGAGTGTGATTAACAAGTCTC 5700
 Db 5641 TACCATTAATTAATGATGACAAACAGTTTATGTCTGAGTGTGATTAACAAGTCTC 5700

Qy 5701 CTTTTCCTCCATTTCTTTTGGGCAATGAGAGCTGAGAGCTTGCTCCCTAAGAG 5760
 Db 5701 CTTTTCCTCCATTTCTTTTGGGCAATGAGAGCTGAGAGCTTGCTCCCTAAGAG 5760
 Qy 5761 AGCTCATGATGATGACACTCACTCTGTATGCTCTCTATCTATCTCCAGAGAGATGATC 5820
 Db 5761 AGCTCATGATGATGACACTCACTCTGTATGCTCTCTATCTATCTCCAGAGAGATGATC 5820
 Qy 5821 TTCTTTCCACTGAGAGAGCTCTGCCCATGTGATTTCTTGGATTTCCAGAGCAACGTGG 5880
 Db 5821 TTCTTTCCACTGAGAGAGCTCTGCCCATGTGATTTCTTGGATTTCCAGAGCAACGTGG 5880
 Qy 5881 CCTGTGATGAGAAAAAGAACTCCGGAATTTGCTTAATTTGAGAGCACTCACTCTA 5940
 Db 5881 CCTGTGATGAGAAAAAGAACTCCGGAATTTGCTTAATTTGAGAGCACTCACTCTA 5940
 Qy 5941 TTTTTCCTTAATTTCAATTTGCTTCTATCTATCTGAGATTTGTTAGGTTAATTTT 6000
 Db 5941 TTTTTCCTTAATTTCAATTTGCTTCTATCTATCTGAGATTTGTTAGGTTAATTTT 6000
 Qy 6001 TTTTTCAGCCCAATTTGATCTGACCTTGAATTTAACTTGAGATCACTCTCTACT 6060
 Db 6001 TTTTTCAGCCCAATTTGATCTGACCTTGAATTTAACTTGAGATCACTCTCTACT 6060
 Qy 6061 TTTTTCAGCCCAATTTGATCTGACCTTGAATTTAACTTGAGATCACTCTCTACT 6120
 Db 6061 TTTTTCAGCCCAATTTGATCTGACCTTGAATTTAACTTGAGATCACTCTCTACT 6120
 Qy 6121 AAAAATGATCAATGATGAG 6180
 Db 6121 AAAAATGATCAATGATGAG 6180
 Qy 6181 CCGGACAAATTAATTTCTCTTTTATCTTATCTTATTTCTTATCTGCAATTCAGGAATTA 6240
 Db 6181 CCGGACAAATTAATTTCTCTTTTATCTTATCTTATTTCTTATCTGCAATTCAGGAATTA 6240
 Qy 6241 TACTAATTTTCTCTCTGAG 6300
 Db 6241 TACTAATTTTCTCTCTGAG 6300
 Qy 6301 TCAATGAGTCTGATTAATTTCTCCAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6360
 Db 6301 TCAATGAGTCTGATTAATTTCTCCAAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6360
 Qy 6361 GCATGATATCTCTCTTTTCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6420
 Db 6361 GCATGATATCTCTCTTTTCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6420
 Qy 6421 CCTCTCATCAATGTTTGCATCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6480
 Db 6421 CCTCTCATCAATGTTTGCATCAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6480
 Qy 6481 AATGAATAGGTTTATCAATTTGTATATCTCAGAGAACTTATAGTATGAGTATGAGTATGAGTAT 6540
 Db 6481 AATGAATAGGTTTATCAATTTGTATATCTCAGAGAACTTATAGTATGAGTATGAGTATGAGTAT 6540
 Qy 6541 TCCCAACTTCAATTTACCTGCTGAG 6600
 Db 6541 TCCCAACTTCAATTTACCTGCTGAG 6600
 Qy 6601 CTGTTTATTTGCTGAG 6660
 Db 6601 CTGTTTATTTGCTGAG 6660
 Qy 6661 ATGCTTTCTAGGGGTTTGGCAAG 6720
 Db 6661 ATGCTTTCTAGGGGTTTGGCAAG 6720
 Qy 6721 CAATTTAGAGTTTGGGCACTCATCAATATTTCTTCAACCAATAGAGTGTGCTCA 6780
 Db 6721 CAATTTAGAGTTTGGGCACTCATCAATATTTCTTCAACCAATAGAGTGTGCTCA 6780

QY	6781	AATGAATTCGATGAAGAAATTTTCTCATCTAAATTAATGATGGTCTCAATGCT	6840
Db	6781	AATGAATTCGATGAAGAAATTTTCTCATCTAAATTAATGATGGTCTCAATGCT	6840
QY	6841	TACATGTGCTTCAATTTTATTTTCCATTCCTCAAAATCTACATTTGCCATTAGGCTT	6900
Db	6841	TACATGTGCTTCAATTTTATTTTCCATTCCTCAAAATCTACATTTGCCATTAGGCTT	6900
QY	6901	CTCATGCAATGCAATTCCTTCATTGAAATGAAGCTTTATGAAAAGCAATTTGCTGCTTATG	6960
Db	6901	CTCATGCAATGCAATTCCTTCATTGAAATGAAGCTTTATGAAAAGCAATTTGCTGCTTATG	6960
QY	6961	GAATAGGCACTAGGAGTAAATAGTAAATAGTGTCTGTCTGCAATGATCTGACACT	7020
Db	6961	GAATAGGCACTAGGAGTAAATAGTAAATAGTGTCTGTCTGCAATGATCTGACACT	7020
QY	7021	GAGTTAATTTCTCACCCACAGAGTCCCGCATTTTTCACACATCTCTAGCGAATCCCATTT	7080
Db	7021	GAGTTAATTTCTCACCCACAGAGTCCCGCATTTTTCACACATCTCTAGCGAATCCCATTT	7080
QY	7081	TCCTCTGGTTCATTAATGCAATCTTTTTCCTGTCCAGATGACACAGTCTGGTCATG	7140
Db	7081	TCCTCTGGTTCATTAATGCAATCTTTTTCCTGTCCAGATGACACAGTCTGGTCATG	7140
QY	7141	AGGGTGTCCACACCACTCTTTGTGTATCTGAAATTCCTCCACCTGAGAGAAATTTTCAAG	7200
Db	7141	AGGGTGTCCACACCACTCTTTGTGTATCTGAAATTCCTCCACCTGAGAGAAATTTTCAAG	7200
QY	7201	CCCAAGATAGAGTAACTCATCGGGGTCCACAGACACTGGCTAGATGAGTGGGGGTGTTTGAT	7260
Db	7201	CCCAAGATAGAGTAACTCATCGGGGTCCACAGACACTGGCTAGATGAGTGGGGGTGTTTGAT	7260
QY	7261	CTTAATGTTATCCCATGTCAGCACAGAACTTGTGTGCACTAGAGAGAGGTCAAGCTTC	7320
Db	7261	CTTAATGTTATCCCATGTCAGCACAGAACTTGTGTGCACTAGAGAGAGGTCAAGCTTC	7320
QY	7321	AGAGTCAACAAGAACTGCAATTTCCAACTGGAATTTGAGGACCCCCACTTTTGAATAGTGA	7380
Db	7321	AGAGTCAACAAGAACTGCAATTTCCAACTGGAATTTGAGGACCCCCACTTTTGAATAGTGA	7380
QY	7381	CTTAATTCCTGCGAGTCTCTGATCTCTCTCTTAAATGAGGACAGTAAATCCCACTAG	7440
Db	7381	CTTAATTCCTGCGAGTCTCTGATCTCTCTCTTAAATGAGGACAGTAAATCCCACTAG	7440
QY	7441	CAGGGTGTGGGGAGAACTCAGAGATCAAAACAGCTGTGATCAACTGTGGTTCTGTTTCC	7500
Db	7441	CAGGGTGTGGGGAGAACTCAGAGATCAAAACAGCTGTGATCAACTGTGGTTCTGTTTCC	7500
QY	7501	AGGGTCAATCAGACTGGGGTTTCTGAGCATGGAATTCACCAATCCCAAGTCTTGGGTACAGA	7560
Db	7501	AGGGTCAATCAGACTGGGGTTTCTGAGCATGGAATTCACCAATCCCAAGTCTTGGGTACAGA	7560
QY	7561	CTGACACCAATCAACGGACGTGAGGAGACTCTTGCTCAAGCAGAACCTGAGCTTCAAG	7620
Db	7561	CTGACACCAATCAACGGACGTGAGGAGACTCTTGCTCAAGCAGAACCTTGAAGCTTCAAG	7620
QY	7621	GGGCTGACGTGATCTGTTTCCCTTGTCCGCGCTGACAGGAAACGGGGTGTGCTCTGAGCTC	7680
Db	7621	GGGCTGACGTGATCTGTTTCCCTTGTCCGCGCTGACAGGAAACGGGGTGTGCTCTGAGCTC	7680
QY	7681	CTGGGCTGCGCATGCGCAGGAAACGCTGTCTTCATCTACATCTTCAACTGAGTGGCGGC	7740
Db	7681	CTGGGCTGCGCATGCGCAGGAAACGCTGTCTTCATCTACATCTTCAACTGAGTGGCGGC	7740
QY	7741	GACTTCCCTTCTCTTAAGGCGCACTTAATATGTTGCGGTTACCGCTCATAAATATCCGC	7800
Db	7741	GACTTCCCTTCTCTTAAGGCGCACTTAATATGTTGCGGTTACCGCTCATAAATATCCGC	7800
QY	7801	CATCCCATCTCCAAATCTCAGTCTGTGATGACCTTCCCTAATTTAATAGGCTTAAGC	7860
Db	7801	CATCCCATCTCCAAATCTCAGTCTGTGATGACCTTCCCTAATTTAATAGGCTTAAGC	7860
QY	7861	ATGCTGAAGCGCCATCAGACCGAGCGTGCCTGTTCATCTGTGTGGCCATCTGTACAC	7920

D	b		7861	ATGCTAGACGGCCATTCAGAACCGAAGGGCTGCCTGTCTCATCCTGTGGCCATCTGGTAAAC	7920
O	y		7921	TGCCGCGCCCCAGATACTTCAATCGGTACTGTGTGTCTCTGTCTGGGCCCCTGTCCCTG	7980
D	b		7921	TGCCGCGCCCCAGATACTTCAATCGGTACTGTGTGTCTCTGTCTGGGCCCCTGTCCCTG	7980
O	y		7991	CTGGCGAGTAATCCCTGGAGTGGAATTCTGTGACCTTCCTGTTTAACTGGAGCGTAATCTGT	8040
D	b		7991	CTGGCGAGTAATCCCTGGAGTGGAATTCTGTGACCTTCCTGTTTAACTGGAGCGTAATCTGT	8040
O	y		8041	TGATGTGAACGTACAGATTTTCATTACAATCGCGGCGTGTTTTTTATATGTGTGTCTC	8100
D	b		8041	TGATGTGAACGTACAGATTTTCATTACAATCGCGGCGTGTTTTTTATATGTGTGTCTC	8100
O	y		8101	TGTGGGTCCAGCCCTGTCTCTGTGTCAGATTTCTGTGTGANTCCGGAGAATGCCGCTG	8160
D	b		8101	TGTGGGTCCAGCCCTGTCTCTGTGTCAGATTTCTGTGTGANTCCGGAGAATGCCGCTG	8160
O	y		8161	ACCAGGCTGTATCGTGAACCATCTCTCTCAAGATGTGTCTTCTCTCTGTGTGGCTGCC	8220
D	b		8161	ACCAGGCTGTATCGTGAACCATCTCTCTCAAGATGTGTCTTCTCTCTGTGTGGCTGCC	8220
O	y		8221	TTTGGCATTGAGTGGGCCCCGTGTTCACAGATCCAGATTCGATTTGAAAAGTCTTATTTGT	8280
D	b		8221	TTTGGCATTGAGTGGGCCCCGTGTTCACAGATCCAGATTCGATTTGAAAAGTCTTATTTGT	8280
O	y		8281	CATGTGATCTTAGTTCATTTTCTGTCTGTCCGCTCTTAAACAGACAGTGGCCAAACCATCAT	8340
D	b		8281	CATGTGATCTTAGTTCATTTTCTGTCTGTCCGCTCTTAAACAGACAGTGGCCAAACCATCAT	8340
O	y		8341	TACTTCTTGTGGGCTCTTTTAGGACAGGTAAATAATGACGAACCTGAGCTGTCTC	8400
D	b		8341	TACTTCTTGTGGGCTCTTTTAGGACAGGTAAATAATGACGAACCTGAGCTGTCTC	8400
O	y		8401	CAGAGGGCTCTGACAGAACGCCCTGAGGTGATTAAGGTGAGGGGTGGCTTCTCTDAGAA	8460
D	b		8401	CAGAGGGCTCTGACAGAACGCCCTGAGGTGATTAAGGTGAGGGGTGGCTTCTCTDAGAA	8460
O	y		8461	ACCCTGAGCTGTGCGGAGACAGATTGAGACAGTGAAGAGAACCTCTGCCCTGTCAAGAC	8520
D	b		8461	ACCCTGAGCTGTGCGGAGACAGATTGAGACAGTGAAGAGAACCTCTGCCCTGTCAAGAC	8520
O	y		8521	AGGACTTTGAGAGCAATCTGCCCTGCCACCTTGAACATTAATGCAATTTTCTTACCC	8580
D	b		8521	AGGACTTTGAGAGCAATCTGCCCTGCCACCTTGAACATTAATGCAATTTTCTTACCC	8580
O	y		8581	TTTGTGCTCAGAAATGTCTCAGGGTCCCGAAGGCCCTTACCA	8622
D	b		8581	TTTGTGCTCAGAAATGTCTCAGGGTCCCGAAGGCCCTTACCA	8622

RESULT 2
AC108007/c 91923 bp DNA linear PRI 03-SEP-20022.

LOCUS	AC108007	Homo sapiens chromosome 11, clone CTD-303B.L12, complete sequence.
DEFINITION	AC108007	AC108007.5 GI:22655822
VERSION	HTG.	Homo sapiens chromosome 11, clone CTD-303B.L12
KEYWORDS	ORGANISM	Homo sapiens (human)
SOURCE	UNPUBLISHED	Britten,B., Linton,L., Nusbaum,C., Lander,E., Allen,A., Brown,A., Cameron,J., Campbell,A., Chang,J., Chazaro,B., Cooke,P., Colangelo,M., Collins,S., Collymore,A., Cook,R., Cowie,P., Dekrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fairo,S.,

Ferreira, P., FitzHugh, W., Gage, D., Galaan, J., Gadda, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karakas, A., Kells, C., Laroque, K., Lamares, R., Landers, T., Lenoczky, J., Levine, R., Liu, G., MacLean, C., MacDonald, P., Major, J., Margolis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Mardix, J., Mensu, L., Milnova, T., Mlenga, V., Murphy, T., Taylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Rette, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schwick, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Teefay, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triglilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zaimoun, J., Zembek, L., Zimmer, A., and Zody, M.

[illegible]

JOURNAL Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome

REFERENCE 3 (bases 1 to 91923)

AUTHORS

Barn, B., Nussbaum, C., Lander, E. A. I., Allen, N., Anderson, S.,
Barin, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalov, B.,
Camara, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Detrellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gargava, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamet, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Ton, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Meneses, L., Mihova, T., Munga, V.,
Murphy, T., Naylor, T., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Plunhthang, P., Pierre, N., Raymond, S., Schnuppach, R., Seman, C.,
Roman, J., Roy, A., Schauer, S., Schnuppach, R., Seman, C., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Teefelge, S., Theodore, T., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission

REFERENCE
JOURNAL
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 91923)

AUTHORS

Barna, N., Basselin, V., Bloom, T., Cocquytchek, L., Boukhalter, B., Camarata, J., Chang, J., Chazot, B., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fairo, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gird, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karakas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, V., McCarthy, M., Meldrum, J., Meneses, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, U., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Punthuang, P., Pierre, N., Raymond, C., Retta, R., Riese, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuppach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zember, L., Zimmer, A. and Zody, M.

TITLE Direct Submission

JOURNAL Submitted (03-SEP-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Sep 3, 2002 this sequence version replaced gi:22474954.

----- Genome Center
p://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBF

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

```
----- Project Information
Center project name: L24596
Center clone name: 3038_L_12
```

Only the first 91.9 kb of this clone are being submitted.
The remainder overlaps accession number AC055860 [WIGR project
L9785].

FEATURES

Location/Qualifiers

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/mag="11"
repeat_region
25..70
/rpt_family="L2"
complement(251..899)
/rpt_family="L2"
900..1001
/rpt_family="L1MD2"
1082..1655
/rpt_family="L1MD2"
1671..1850
/rpt_family="L1MD3"
complement(1658..1917)
/rpt_family="L2"
1918..2273
/rpt_family="THE1B"
complement(2274..2327)
/rpt_family="L2"
4505..4713
/rpt_family="L2"
5536..5795
/rpt_family="A1uSc"
5796..5837
/rpt_family="TAAA)n"
5838..5875
complement(7186..7526)
/rpt_family="L1MB4"
complement(10716..10825)
/rpt_family="MIR"
11369..11466
/rpt_family="L2"
complement(11479..11644)
/rpt_family="MIR3"
12072..12409
/rpt_family="MER2"
12692..12875
/rpt_family="L1MD3"
complement(12876..13294)
/rpt_family="MST1"
13395..13552
/rpt_family="L1MD3"
complement(14619..14992)
/rpt_family="L1MB1"
complement(15002..17751)
/rpt_family="L1MB1"
17752..18590
/rpt_family="L1PA6"
19591..20184
/rpt_family="L1PA4"
20185..24476
/rpt_family="L1PA6"
24611..24637
/rpt_family="(T)n"
complement(24639..26213)
/rpt_family="SVA"
26219..26361
/rpt_family="L1MB5"

```



```

repeat_region 26372..26560
/rpc_family="PLAM_C"
repeat_region complement (26636..26719)
/rpc_family="MIR"
repeat_region complement (26863..27139)
/rpc_family="MER41C"
repeat_region complement (27140..27449)
/rpc_family="Alusq"
repeat_region complement (27450..27494)
/rpc_family="MER41C"
repeat_region complement (27495..27796)
/rpc_family="Alusg"
repeat_region complement (27797..28030)
/rpc_family="MER41C"
repeat_region complement (28194..28287)
/rpc_family="MIR"
repeat_region complement (29002..29304)
/rpc_family="Alusd"
repeat_region complement (29337..29755)
/rpc_family="MSTa"
repeat_region 30412..30452
/rpc_family="(TG)n"
repeat_region complement (31075..31197)
/rpc_family="MIR"
repeat_region 31198..31246
/rpc_family="(CA)n"
repeat_region complement (31247..31341)
/rpc_family="MIR"
repeat_region 31541..31821
/rpc_family="Alusx"
repeat_region 31822..31850
/rpc_family="(GA)n"
repeat_region 32030..32230
/rpc_family="MER5A"
repeat_region complement (32650..32747)

```

Query Match 99.5%; Score 8574.6; DB 9; Length 91923;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 8594; Conservative 0; Mismatches 14; Indels 1; Gaps 1;

```

QY 1 TGTATGAGCAATGTCACTTAATACCAAAACGAGAAAGATATCAAAAAAGAAAC 60
DB 17003 TGTATGAGCAATGTCACTTAATACCAAAACGAGAAAGATATCAAAAAAGAAAC 16544
QY 61 TATAGACCACTACCTGATGAATATACATGCAAGAAATCCCAAAATATCTAGCTAAC 120
DB 16943 TATAGACCACTACCTGATGAATATACATGCAAGAAATCCCAAAATATCTAGCTAAC 16884
QY 121 CCAATCCACAGCATATCAAGAAATATCCACATTTGTCAAGTGGTTTCATACCAAGG 180
DB 16883 TCAATCCACAGCATATCAAGAAATATCCACATTTGTCAAGTGGTTTCATACCAAGG 16824
QY 181 GTGAGAGATAGTTTAAATACACAAAGTCAATTAATGTGATCATCACATTAACGAATTA 240
DB 16823 GTGAGAGATAGTTTAAATACACAAAGTCAATTAATGTGATCATCACATTAACGAATTA 16764
QY 241 AAAACAAAAATCAATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAATCTAAC 300
DB 16763 AAAACAAAAATCAATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAATCTAAC 16704
QY 301 ATTTCTTTATGATTTAAACCTTCACGAAATCGACATAGAAAGACATACCTTAATGTAA 360
DB 16703 ATTTCTTTATGATTTAAACCTTCACGAAATCGACATAGAAAGACATACCTTAATGTAA 16644
QY 361 TAAAGCCATATATGAGGAGCCACAGCAAACTTATATGAAATGGGAAAAAGTTGAAAA 420
DB 16643 TAAAGCCATATATGAGGAGCCACAGCAAACTTATATGAAATGGGAAAAAGTTGAAAA 16584
QY 421 CATTTGCTCTGAGAACTGGAACAAGAGATCTATTCAACCTTCTATTCAACAT 480
DB 16583 CATTTGCTCTGAGAACTGGAACAAGAGATCTATTCAACCTTCTATTCAACAT 16524
QY 481 AGTAGTGAAGTTTACGCAAGCAATCAGCAAGAAAGAAATCAAGGGCAACCAAT 540

```

```

DB 16523 AGTAGTGAAGTTTACGCAAGCAATCAGCAAGAAAGAAATCAAGGGCAACCAAT 16464
QY 541 CAATTAAGAGAAAGTCAAACTGTCCCTGTTCACTGATATATGATTTATACCTAGAAA 600
DB 16463 CAATTAAGAGAAAGTCAAACTGTCCCTGTTCACTGATATATGATTTATACCTAGAAA 16404
QY 601 CCTTAAGAACTCATCCAGAAAGCTCCTAGAACTGATCATTAATTCAGTAAAGTTTCAAG 660
DB 16403 CCTTAAGAACTCATCCAGAAAGCTCCTAGAACTGATCATTAATTCAGTAAAGTTTCAAG 16344
QY 661 ATTCAACTAAATGTATACAAATCAGTAGCACTGTATACCAAGCAAGTCAAGCTGA 720
DB 16343 ATTCAACTAAATGTATACAAATCAGTAGCACTGTATACCAAGCAAGTCAAGCTGA 16284
QY 721 GAATCAATCAAGAACTCAAACTTTTAAATAGCTGTAAAAAATACTTAAGATAAT 780
DB 16283 GAATCAATCAAGAACTCAAACTTTTAAATAGCTGTAAAAAATACTTAAGATAAT 16224
QY 781 CTTAACCAGAGGTGAAGACCTCTACAGAGAAACTTACAAAACAGCTGACATCATA 840
DB 16223 CTTAACCAGAGGTGAAGACCTCTACAGAGAAACTTACAAAACAGCTGACATCATA 16164
QY 841 GATGACCAAAACAAGTGAACAATCCCAATGCTCATGAGTGGGTAGAAATCAATTTGTG 900
DB 16163 GATGACCAAAACAAGTGAACAATCCCAATGCTCATGAGTGGGTAGAAATCAATTTGTG 16104
QY 901 AAATGACATATTTGCCAAAGCAATCTACAAAGTTCAATCCCAACCAATATCA 960
DB 16103 AAATGACATATTTGCCAAAGCAATCTACAAAGTTCAATCCCAACCAATATCA 16044
QY 961 TCATCATTTCTTCACAGAACTGAAAAACAATTTCAAAATTCATATGAAACAACA 1020
DB 16043 TCATCATTTCTTCACAGAACTGAAAAACAATTTCAAAATTCATATGAAACAACA 15984
QY 1021 AAAAAAAAAAAAAAAAAACCCGATAGCCAAAGCAAGCTTGACAAAAGAAACAAATCTGGA 1080
DB 15983 AAAAAAAAAAAAAAAAAACCCGATAGCCAAAGCAAGCTTGACAAAAGAAACAAATCTGGA 15924
QY 1081 GGCATCACTTACCATCTTCAAACTATCTACAAAGCTTAATCAACCAAAACATCATGG 1140
DB 15923 GGCATCACTTACCATCTTCAAACTATCTACAAAGCTTAATCAACCAAAACATCATGG 15864
QY 1141 CACTGACATTAACCTAGGACATAGCAATAGAAAAAGAGAAATCCAGAAATTAAG 1200
DB 15863 CACTGACATTAACCTAGGACATAGCAATAGAAAAAGAGAAATCCAGAAATTAAG 15804
QY 1201 CCAATTAATTTAGCCCACTGATTTTGAACAAGCAACAAAACATTAAGTGGGAAAA 1260
DB 15803 CCAATTAATTTAGCCCACTGATTTTGAACAAGCAACAAAACATTAAGTGGGAAAA 15744
QY 1261 GACATTTAGTTAAACAATGCTGAGATTAATGCAACCAATGATGAAAGATGAAA 1320
DB 15743 GACATTTAGTTAAACAATGCTGAGATTAATGCAACCAATGATGAAAGATGAAA 15684
QY 1321 CTGATCCCTTGTCTCACTTAATCAAAAATTTGATACAAAGTATCAAGAATTAA 1380
DB 15683 CTGATCCCTTGTCTCACTTAATCAAAAATTTGATACAAAGTATCAAGAATTAA 15625
QY 1381 TCTGAGACCTTAACCAATTAATTTCTAGAAATTAATCAAGAAATGCTTCTAGACAT 1440
DB 15624 TCTGAGACCTTAACCAATTAATTTCTAGAAATTAATCAAGAAATGCTTCTAGACAT 15565
QY 1441 TCACTTAGGCAAAAGACTTCAATGCGCAAGAACCCAAAGTAAATGCACAAAAACAAAAT 1500
DB 15564 TCACTTAGGCAAAAGACTTCAATGCGCAAGAACCCAAAGTAAATGCACAAAAACAAAAT 15505
QY 1501 AAATGATAGGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1560
DB 15504 AAATGATAGGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 15445
QY 1561 AGCAAAAGCAACCCACGAGTGAAGAAATCTTCAAACTAAGCATCTGACTAAG 1620

```

Db	15444	AGCAAA	CAGACA	CAACCC	ACCGAG	TGAGAG	AAAAATCTT	CA	CAAACT	TAAGATCT	GACCAAG	15385
OY	1621	ACTAATAT	CCGGAAT	CCACAAG	GAACCT	CAAA	CAAT	CAGAGA	GAAGA	CAAA	CAATCC	1680
Db	15384	ACTAATAT	CCGGAAT	CCACAAG	GAACCT	CAAA	CAAT	CAGAGA	GAAGA	CAAA	CAATCC	15325
OY	1681	CATTGA	AGATGGG	CTTAAG	CACTGA	ATTG	CA	AAAGAA	GGATTTC	CAATGGC	1740	
Db	15324	CATGAA	AGATGGG	CTAAG	CACTGA	ATTG	CA	AAAGAA	GGATTTC	CAATGGC	15265	
OY	1741	CAACAA	CAGAAAAA	ATGCTTAA	CATCACTA	TAAGATTA	GGAAATG	TAATCA	CACTG	1800		
Db	15264	CAACAA	CAGAAAAA	ATGCTTAA	CATCACTA	TAAGATTA	GGAAATG	TAATCA	CACTG	15205		
OY	1801	TAATGGA	TACCACTT	ACTCTG	CCAGAA	TGTCAT	TAATTTAAAA	CTAAAA	TAAATA	1860		
Db	15204	TAATGGA	TACCACTT	ACTCTG	CCAGAA	TGTCAT	TAATTTAAAA	CTAAAA	TAAATA	15145		
OY	1861	GATGTTGG	TGGGTCTGTG	TGATTA	AGGAACA	CTTTT	CACTGCTGG	TGGGAATGAAC	1920			
Db	15144	GATGTTGG	TGGGTCTGTG	TGATTA	AGGAACA	CTTTT	CACTGCTGG	TGGGAATGAAC	15085			
OY	1921	TTGGCG	CAACCACT	ATGGA	AAAA	CAGTGG	AAATTTCTTA	AGAACTAA	AGTAGATCGAC	1980		
Db	15084	TTGGCG	CAACCACT	ATGGA	AAAA	CAGTGG	AAATTTCTTA	AGAACTAA	AGTAGATCGAC	15025		
OY	1981	CATTGAT	CCAGCAATCC	CAATTA	ATAATG	TATAATAT	TTATTTATTA	TATACATGAAT	2040			
Db	15024	CATTGAT	CCAGCAATCC	CAATTA	ATAATG	TATAATAT	TTATTTATTA	TATACATGAAT	14965			
OY	2041	ACAACT	CAGCAT	TAATAA	AGAAATTA	ATGATGA	CATTCA	CAGCAATCT	TAAGATGG	2100		
Db	14964	ACAACT	CAGCAT	TAATAA	AGAAATTA	ATGATGA	CATTCA	CAGCAATCT	TAAGATGG	14905		
OY	2101	AGACCC	TTATTTCTA	AGTGGG	TAACTCA	GAATGGA	AAAA	CAATCATATGTTCT	CAC	2160		
Db	14904	AGACCC	TTATTTCTA	AGTGGG	TAACTCA	GAATGGA	AAAA	CAATCATATGTTCT	CAC	14855		
OY	2161	TTACA	GTTGGGGCTA	GTGTG	GAACA	CGAAG	CA	TAGAATGAT	TATGAACTCTGG	2220		
Db	14844	TTACA	GTTGGGGCTA	GTGTG	GAACA	CGAAG	CA	TAGAATGAT	TATGAACTCTGG	14785		
OY	2221	GGA	CTTGAGGG	GAAGATG	GAAGA	GAGCG	GAATAA	AGA	CTACATG	GGTACGT	2280	
Db	14784	GGA	CTTGAGGG	GAAGATG	GAAGA	GAGCG	GAATAA	AGA	CTACATG	GGTACGT	14725	
OY	2281	GTACA	CTGCTCAG	GTGATGG	TGACCA	AAATCT	CAGAA	TTTACA	CTA	TAAGAACTTATC	2340	
Db	14724	GTACA	CTGCTCAG	GTGATGG	TGACCA	AAATCT	CAGAA	TTTACA	CTA	TAAGAACTTATC	14655	
OY	2341	CATGA	AGCAAA	CA	CACTGTTCC	CCAAAA	TCCAA	TGAATA	TAATAAT	TAATA	2400	
Db	14664	CATGA	AGCAAA	CA	CACTGTTCC	CCAAAA	TCCAA	TGAATA	TAATAAT	TAATA	14605	
OY	2401	ATGATTT	AATTTCA	CAGAA	TTTAAAA	AGTTCA	CTGTTCA	AGTTTAT	TAATGA	AGTA	2460	
Db	14604	ATGATTT	AATTTCA	CAGAA	TTTAAAA	AGTTCA	CTGTTCA	AGTTTAT	TAATGA	AGTA	14545	
OY	2461	AGAA	TAAAAAG	TGAAG	AGTGA	AGCTCTG	GA	CAAGGA	CTTAATTT	CACTTG	2520	
Db	14544	AGAA	TAAAAAG	TGAAG	AGTGA	AGCTCTG	GA	CAAGGA	CTTAATTT	CACTTG	14485	
OY	2521	CATACA	CTTCTCG	GCATTTG	GAAGA	AGATACA	CACTGA	TAATAT	CA	CACTATGA	2580	
Db	14484	CATACA	CTTCTCG	GCATTTG	GAAGA	AGATACA	CACTGA	TAATAT	CA	CACTATGA	14425	
OY	2581	AAGAAA	CA	CAAAAA	TTGTG	CAAG	CACTG	TGATG	ATGTC	CCATGATCTAC	2640	
Db	14424	AAGAAA	CA	CAAAAA	TTGTG	CAAG	CACTG	TGATG	ATGTC	CCATGATCTAC	14365	
OY	2641	TTTTT	CAATTTG	GAATATA	AGCTCAT	ATATG	ATGCA	CTTGCTTA	TATGATG	CTG	2700	
Db	14364	TTTTT	CAATTTG	GAATATA	AGCTCAT	ATATG	ATGCA	CTTGCTTA	TATGATG	CTG	14305	

QY	2701	GGCCATAATGATTTTAAATAAGGACAGCAATTACTTAACAGGAAGTTGAACAAGATGG	2760
Db	14304	GCCCATATATGATTTTAAATAGGACAGCAATTACTTAACAGGAAGTTGAACAAGATGG	14245
QY	2761	GACGTGATAGAGAGGCTTAAATGTACTGSAATAGGGAAGAAGCCAAAGATCATCTGAG	2820
Db	14244	GACGTGATAGAGAGGCTTAAATGTACTGSAATAGGGAAGAAGCCAAAGATCATCTGAG	14185
QY	2821	TTAGGATTTTGTCTCAAAATACCTCTGGCTCTGATTTGGCCATAGTCTCTCATAGGA	2880
Db	14184	TTAGGATTTTGTCTCAAAATACCTCTGGCTCTGATTTGGCCATAGTCTCTCATAGGA	14125
QY	2881	ATAACAGACTGTCCAGCATCTTCGTAACCTGSAATGTCTCAACAGCTTCATTCACGT	2940
Db	14124	ATAACAGACTGTCTCAAGATCTTCGTAACCTGSAATGTCTCAACAGCTTCATTCACGT	14065
QY	2941	CTGTAGGCAATCTCTCGAATTAAGCAACAAGAAAAGTCTCTGAAGTCACTGAATCCCA	3000
Db	14064	CTGTAGGCAATCTCTCGAATTAAGCAACAAGAAAAGTCTCTGAAGTCACTGAATCCCA	14005
QY	3001	GAAGGCTCTCTACCTTTAGCAACAAGGAGGCTTCAACCACTGGACAAGAAAGAAAGAT	3060
Db	14004	TAAAGCTCTCTACCTTTAGCAACAAGGAGGCTTCAACCACTGGACAAGAAAGAAAGAT	13945
QY	3061	AAGGTAGTACCAAGAACTCTCTTCTTCCACAGTCAAGTTATGATTTTGTCTGAAGATC	3120
Db	13944	AAGGTAGTACCAAGAACTCTCTTCTTCCACAGTCAAGTTATGATTTTGTCTGAAGATC	13885
QY	3121	ATGTCTTATGCTTCCACCTTGGGTGTACATGAGGGGGTCAAGAGCTGTTCAGGAAA	3180
Db	13884	ATGTCTTATGCTTCCACCTTGGGTGTACATGAGGGGGTCAAGAGCTGTGTTCAGGAAA	13825
QY	3181	AGACAGAGACATGAAGCTTCTTTTCAAGAACTGAGTGTGTCAACCCAACTGTGTGAG	3240
Db	13824	AGACAGAGACATGAAGCTTCTTTTCAAGAACTGAGTGTGTCAACCCAACTGTGTGAG	13765
QY	3241	CTCTAAATGGTGTCCCCCTTCTAATTTATCTCCCATATCACCTCTTCAATTCGAATCA	3300
Db	13764	CTCTAAATGGTGTCCCCCTTCTAATTTATCTCCCATATCACCTCTTCAATTCGAATCA	13705
QY	3301	TTCAATCTGCCCTCATGGAAGACCTGTGCTCTTACATTCATTTTAAAGAGAAAGGAC	3360
Db	13704	TTCAATCTGCCCTCATGGAAGACCTGTGCTCTTACATTCATTTTAAAGAGAAAGGAC	13645
QY	3361	ATGCAAGCATTTCTTCCAGAGTTGAACGTGTATAGAGCCAGTTCTTTGTTCACTTAC	3420
Db	13644	ATGCAAGCATTTCTTCCAGAGTTGAACGTGTATAGAGCCAGTTCTTTGTTCACTTAC	13585
QY	3421	TTTTCAAAATTTATCTCTTCTTGTGCTATCTGGAAGAGCTTAAGGAAGATATATAGGCCCA	3480
Db	13584	TTTTCAAAATTTATCTCTTCTTGTGCTATCTGGAAGAGCTTAAGGAAGATATATAGGCCCA	13525
QY	3481	ATAATTAAGAGTGTTCATAGAGAAAGTATTTCAAAAGATGACACAGATTAAAGGTGAG	3540
Db	13524	ATAATTAAGAGTGTTCATAGAGAAAGTATTTCAAAAGATGACACAGATTAAAGGTGAG	13465
QY	3541	GATCTTAAGCAATTCATAGGGGAGCACTACTTCTCTCCCTAGGCTGAACCGACAGG	3600
Db	13464	GATCTTAAGCAATTCATAGGGGAGCACTACTTCTCTCCCTAGGCTGAACCGACAGG	13405
QY	3601	GAAGGACAGATACCATGTGTGGCATAGCAATAGCTGTAGCCATAAGGTTGGGAAGAT	3660
Db	13404	GAAGGACAGATACCATGTGTGGCATAGCAATAGCTGTAGCCATAAGGTTGGGAAGAT	13345
QY	3661	GAGCAGGCAAGTGAAGAAAGCTGTGTGGCAACGACAGCCACACAGGCTGATATAGTT	3720
Db	13344	GAGCAGGCAAGTGAAGAAAGCTGTGTGGCAACGACAGCCACACAGGCTGATATAGTT	13285
QY	3721	TGATATCTGTGTCCCAACCAAAATCTCATGTATGTAAATTTCCAAATGTTGAGGAAGGG	3780
Db	13284	TGATATCTGTGTCCCAACCAAAATCTCATGTATGTAAATTTCCAAATGTTGAGGAAGGG	13225

OY	3781	CTGTGGGAGATGATTAATAGAACAGGGGATGGTTTGCAATGAATGTTTAAACCAAT	3840
Db	13224	CTTTGGGAGATTAATTTAGATCAGGGGATGGTTTGCAATGAATGTTTAAACCAAT	13165
OY	3841	CCCCCTTGGTATTTGTTGTGTGATCTGACGAGTTCTCATGAAATCTAGTTGTTTAAA	3900
Db	13164	CCCCCTTGGTATTTGTTGTGTGATCTGACGAGTTCTCATGAAATCTAGTTGTTTAAA	13105
OY	3901	GCGTGTAGCACTCTCTCCCTCTCTCTTAATCTCTGCTCTCAACATGTGACGCTCTGCTC	3960
Db	13104	GCGTGTAGCACTCTCTCCCTCTCTCTTAATCTCTGCTCTCAACATGTGACGCTCTGCTC	13045
OY	3961	CCCCCTTGGCTTTCACACGAGATTGGAAAGCTTCTGAGAGCCCTCCCGAAGACAAACCTGC	4020
Db	13044	CCCCCTTGGCTTTCACACGAGATTGGAAAGCTTCTGAGAGCCCTCCCGAAGACAAACCTGC	12985
OY	4021	TATGCTTCTGTGACGCTGTGAGAGCTATTAGCCAGTTAAACCATTTCTTCATATAAT	4080
Db	12984	TATGCTTCTGTGACGCTGTGAGAGCTATTAGCCAGTTAAACCATTTCTTCATATAAT	12925
OY	4081	TCCCACTCTCAGTAATTTCTTTTTAGCAATTTGAGAAATGAACTAATACACAGACAGAG	4140
Db	12924	TCCCACTCTCAGTAATTTCTTTTTAGCAATTTGAGAAATGAACTAATACACAGACAGAG	12865
OY	4141	CCAGAGATGGAATATCCCAAGTGCTTTCCTGCTGTCTTCACTCTCTGCTGTCTC	4200
Db	12864	CCAGAGATGGAATATCCCAAGTGCTTTCCTGCTGTCTTCACTCTCTGCTGTCTC	12805
OY	4201	CCAGTGTCTCAATTTCCACAGAAACAGAAATATAAAAGATCCCATGATGTGTACAT	4260
Db	12804	CCAGTGTCTCAATTTCCACAGAAACAGAAATATAAAAGATCCCATGATGTGTACAT	12745
OY	4261	GAAGCCACTCTCTTGGGATGTCAAAACAGATAAAGAAATGAAAGCAAAATCTCATGG	4320
Db	12744	GAAGCCACTCTCTTGGGATGTCAAAACAGATAAAGAAATGAAAGCAAAATCTCATGG	12685
OY	4321	TAAATGAGACTATCCCTCTCACTTCTTGTATCTCTTAATCTGCGGGCTTTCTATTC	4380
Db	12684	TAAATGAGACTATCCCTCTCACTTCTTGTATCTCTTAATCTGCGGGCTTTCTATTC	12625
OY	4381	TGATTTGATCCCTGTCTCATTTCAAGCTCTATACAGCTATTAATGTTGGCTGTCTTTC	4440
Db	12624	TGATTTGATCCCTGTCTCATTTCAAGCTCTATACAGCTATTAATGTTGGCTGTCTTTC	12565
OY	4441	TCTACTGTCACTTTTATGACAGAAATGTTTGCATTTGTTAAAAATGCAAGAAATATAAT	4500
Db	12564	TCTACTGTCACTTTTATGACAGAAATGTTTGCATTTGTTAAAAATGCAAGAAATATAAT	12505
OY	4501	GTAATTTTAAAAAGACATATGTAATTTTGTGTAGAAATATAAGTTTGGCTGATCTAATAA	4560
Db	12504	GTAATTTTAAAAAGACATATGTAATTTTGTGTAGAAATATAAGTTTGGCTGATCTAATAA	12445
OY	4561	GACATGGAAGAAATATCTTAAACAGAAAGTATAGTTGTGCTCTGCGTCACTAGTT	4620
Db	12444	GACATGGAAGAAATATCTTAAACAGAAAGTATAGTTGTGCTCTGCGTCACTAGTT	12385
OY	4621	CTGATATCTACAGATTCAACAAACCTACAGAGAAACCTTTCCAAAAATAAAGGTGGCG	4680
Db	12384	CTGATATCTACAGATTCAACAAACCTACAGAGAAACCTTTCCAAAAATAAAGGTGGCG	12325
OY	4681	GAGTTGTGTATGTACTGAAACAGTACAAACTTGTATTTCTTGTCACTTATTTCTGAAAA	4740
Db	12324	GAGTTGTGTATGTACTGAAACAGTACAAACTTGTATTTCTTGTCACTTATTTCTGAAAA	12265
OY	4741	CTAACAATATAAGAACTTATATAGATTTGCAATTTTGTGCACTTATTTCTTAATTACTTT	4800
Db	12264	CTAACAATATAAGAACTTATATAGATTTGCAATTTTGTGCACTTATTTCTTAATTACTTT	12205
OY	4801	AAATGATTTAATGATCTGGAGAAAGTGCATAGAGTATATAAATAATCACTATATAGG	4860
Db	12204	AAATGATTTAATGATCTGGAGAAAGTGCATAGAGTATATAAATAATCACTATATAGG	12145
OY	4861	AAATTTGAGCATCTGCAGATTTTGTGCTGTGCGGGTCTGAAAAAGATCCCTGTAAAT	4920

Db	12144	AAATTGACATCTGCAGATTTTGGTCTGTGCTGGGTTCTGGAAAGATCCCTGTAAAT	12088
Qy	4921	ACACAAAATGACACTCTTCGAGATCTGAACTAGAAAGCTCCAAAGCATCATATCAGAA	49880
Db	12084	ACACAAAATGACACTCTTCGAGATCTGAACTAGAAAGCTCCAAAGCATCATATCAGAA	12022
Qy	4981	TTCCAAAATTTGCTGCTCCCAAGTTCTAGAGATGTGCTCATCTTTGTATCTTACAT	50404
Db	12024	TTCCAAAATTTGCTGCTCCCAAGTTCTAGAGATGTGCTCATCTTTGTATCTTACAT	11962
Qy	5041	GGTTCACGAGCATTTAGCATTTCCAGTCTTATGAAAAAGGACGAGGGGAAGAGGCT	51002
Db	11964	GGTTCACGAGCATTTAGCATTTCCAGTCTTATGAAAAAGGACGAGGGGAAGAGGCT	11902
Qy	5101	TTGCTCTCTTATTAAATCCCATGAGCCAGACCTTGCTCTGTCACTTTTGTATCTTCC	51602
Db	11904	TTGCTCTCTTATTAAATCCCATGAGCCAGACCTTGCTCTGTCACTTTTGTATCTTCC	11878
Qy	5161	ACTTAAACGACCTGCTCATGGAGTGTCAATCAGCATCAGAGAAAATCTGGATGTGGCT	52202
Db	11884	ACTTAAACGACCTGCTCATGGAGTGTCAATCAGCATCAGAGAAAATCTGGATGTGGCT	11782
Qy	5221	CTTGCTGCTGTTGATCATTTCTCAGAAAAGTTATGTGACCAAAAAGGAAATCTTGGGGC	52802
Db	11784	CTTGCTGCTGTTGATCATTTCTCAGAAAAGTTATGTGACCAAAAAGGAAATCTTGGGGC	11722
Qy	5281	ACCAGCAGTCTCTCAGCCCTGACGTGCTGTGATCTGTGCTCAGATCAAGATTTTTCA	53404
Db	11724	ACCAGCAGTCTCTCAGCCCTGACGTGCTGTGATCTGTGCTCAGATCAAGATTTTTCA	11662
Qy	5341	GGAACTCTCGAATAATAATAATGATGGGGCAGAGAACAGACTGAGTCTGTGCAGAA	54002
Db	11664	GGAACTCTCGAATAATAATAATGATGGGGCAGAGAACAGACTGAGTCTGTGCAGAA	11602
Qy	5401	CTCCAGGAGCAAGGGGCTGTATTTGAGACTGTCTCTTCAATGTTGTGAACACAGAAAACCT	54602
Db	11604	CTCCAGGAGCAAGGGGCTGTATTTGAGACTGTCTCTTCAATGTTGTGAACACAGAAAACCT	11542
Qy	5461	TTAAATCTCTGAGGCTTAGCTTCATCTTATGTTATGAGGATTAATACATAGACAGTCT	55202
Db	11544	TTAAATCTCTGAGGCTTAGCTTCATCTTATGTTATGAGGATTAATACATAGACAGTCT	11482
Qy	5521	TTAAAGAACATCATAGCATGTTTAAACAAACATGCTAAATGTTGTGTATACCAAGTGAAG	55802
Db	11484	TTAAAGAACATCATAGCATGTTTAAACAAACATGCTAAATGTTGTGTATACCAAGTGAAG	11422
Qy	5581	AGACAGCATGACTTACTCTTACGGAATCTTGGGTTTCAATGAGAAAGCAACATATCA	56402
Db	11424	AGACAGCATGACTTACTCTTACGGAATCTTGGGTTTCAATGAGAAAGCAACATATCA	11362
Qy	5641	TACCAATACATTAATGAGCAAAACAGTTTATGTGCTCGAGTGTGATTAACAGAGTTCTC	57002
Db	11364	TACCAATACATTAATGAGCAAAACAGTTTATGTGCTCGAGTGTGATTAACAGAGTTCTC	11302
Qy	5701	CTTTTCCCTCCATTTCTTTTGGGCAATCAGAGCTGTGGCACTTGTCTCCCTAAGAG	57602
Db	11304	CTTTTCCCTCCATTTCTTTTGGGCAATCAGAGCTGTGGCACTTGTCTCCCTAAGAG	11242
Qy	5761	AGCTCATGATGAGTCACTCACTCTGATGTCTCTTAATCTCCCAAGAGAGATGATC	58202
Db	11244	AGCTCATGATGAGTCACTCACTCTGATGTCTCTTAATCTCCCAAGAGAGATGATC	11182
Qy	5821	TTCTTTTCCACTGTGAGAGCTCTGCGCAATGTGCATTTCTGGGATTTCCAGAGCAAACTGG	58802
Db	11184	TTCTTTTCCACTGTGAGAGCTCTGCGCAATGTGCATTTCTGGGATTTCCAGAGCAAACTGG	11122
Qy	5881	CCTCTGATAGGCAAAAAGAACTCTGAATTTGTTCTTAAATGAGCAGCACTCACTCTA	59402
Db	11124	CCTCTGATAGGCAAAAAGAACTCTGAATTTGTTCTTAAATGAGCAGCACTCACTCTA	11062
Qy	5941	TTTTTCCCTTAATTTCAATTTGCTTCTCAATCTATCTGAGAGTTGTTAGGTTAATTTT	60002

Db	11064	TTTTTCCCTATTCTAATTGGCTCTCATCTCTAATCTGAATCTGAAGATTGGTTAGGTTAATTTTT	11005
Qy	6001	TTTTTCAGCCCAATATTTTGACTGTCAACTTGGATTTAATTAAGATACTCTCTACT	6060
Db	11004	TTTTTCAGCCCAATATTTTGACTGTCAACTTGGATTTAATTAAGATACTCTCTACT	10945
Qy	6061	TTAACCCCTCTAATCATGTATAATGCAACAATAAGGGGCTGGGCTCAAAGGGGCTGGGA	6120
Db	10944	TTAACCCCTCTAATCATGTATAATGCAACAATAAGGGGCTGGGCTCAAAGGGGCTGGGA	10885
Qy	6121	AAAAATGATCATGAGTCAGACCCCTGCTGGGCTCAATTATATATATATAATTAACCC	6180
Db	10884	AAAAATGATCATGAGTCAGACCCCTGCTGGGCTCAATTATATATATATAATTAACCC	10825
Qy	6181	CCCGGCAAAATATATCTCTCTTTAATCTTAATTTGATATCTGCAATACAGAAATTA	6240
Db	10824	CCCGGCAAAATATATCTCTCTTTAATCTTAATTTGATATCTGCAATACAGAAATTA	10765
Qy	6241	TACTAATTTTTTACTCTCTAGGCTCTGAGATTAATAAAGGGCAATCTTAATAACG	6300
Db	10764	TACTAATTTTTTACTCTCTAGGCTCTGAGATTAATAAAGGGCAATCTTAATAACG	10705
Qy	6301	TCAATCAGCTGCTGTTATTTCTCCCAAAATTAGACCTAATCTCATCTCCAGTTGAAATTT	6360
Db	10704	TCAATCAGCTGCTGTTATTTCTCCCAAAATTAGACCTAATCTCATCTCCAGTTGAAATTT	10645
Qy	6361	GCATGAATATCTGCTTTTACAACCCGAAGCCCTYACATTTCTCTAATTTCCATCATGAGCT	6420
Db	10644	GCATGAATATCTCTCTTTACAACCCGAAGCCCTYACATTTCTCTAATTTCCATCATGAGCT	10585
Qy	6421	CCTCTCATACAAATGTTTGCATGAAAGAAAGGCTACAAAGATCTCCGGAAGAGAG	6480
Db	10584	CCTCTCATACAAATGTTTGCATGAAAGAAAGGCTACAAAGATCTCCGGAAGAGAG	10525
Qy	6481	AATGAATATAGGTTTACATTTGATTAATCTCAGACAGAACATTAGTCCCCATACATAT	6540
Db	10524	AATGAATATAGGTTTACATTTGATTAATCTCAGACAGAACATTAGTCCCCATACATAT	10465
Qy	6541	TCCGACATTTCAATTACCTGCTGAGTGGCACTCAGGCTCACCTCATCTCTTTCT	6600
Db	10464	TCCGACATTTCAATTACCTGCTGAGTGGCACTCAGGCTCACCTCATCTCTTTCT	10405
Qy	6601	CTGTTCTATTGCTGAGCAATTCACTCAGACCCACACCTTACCCAAACATGTGTACAA	6660
Db	10404	CTGTTCTATTGCTGAGCAATTCACTCAGACCCACACCTTACCCAAACATGTGTACAA	10345
Qy	6661	ATGCTTCTAGGGGTTGGCAAAAGCAACATGAGTCTTATTTTAAAGGCAATCAGTGT	6720
Db	10344	ATGCTTCTAGGGGTTGGCAAAAGCAACATGAGTCTTATTTTAAAGGCAATCAGTGT	10285
Qy	6721	CAATTTCAAGTTTGGGCACTCATCAATCATTTCTTCAACAACAGATAGAGGTGTCACA	6780
Db	10284	CAATTTCAAGTTTGGGCACTCATCAATCATTTCTTCAACAACAGATAGAGGTGTCACA	10225
Qy	6781	AATAGAATTTCTGATGAATGAAATTTTCTTCATCTAATTAATGTGTGTCTTAATGCT	6840
Db	10224	AATAGAATTTCTGATGAATGAAATTTTCTTCATCTAATTAATGTGTGTCTTAATGCT	10165
Qy	6841	TAAATGTGCTTCATTTTATTTTATTTTCAATTCATCCAAATCTAACATGCGATTAAGGCT	6900
Db	10164	TAAATGTGCTTCATTTTATTTTATTTTCAATTCATCCAAATCTAACATGCGATTAAGGCT	10105
Qy	6901	CTCATGATCATCTCTTCATTTGATGAAGCTTTATGAAAAGCAATGTGTGCTTTATG	6960
Db	10104	CTCATGATCATCTCTTCATTTGATGAAGCTTTATGAAAAGCAATGTGTGCTTTATG	10045
Qy	6961	GAATAGGCACTGAGAGTATAAATGTAAATGTGCTCTGTCTGCAATGACATGACACT	7020
Db	10044	GAATAGGCACTGAGAGTATAAATGTAAATGTGCTCTGTCTGCAATGACATGACACT	9985
Qy	7021	GAGTTATTTCTACCCACAGAGTCCCGCATTTTTCACACATCTCTGAGGAATGCCATTT	7080
Db	9984	GAGTTATTTCTACCCACAGAGTCCCGCATTTTTCACACATCTCTGAGGAATGCCATTT	9925

OY	7081	CCCTCGTTCATATATGCAATGATTCCTTTTTCCTGTCAGAGATGACCAAGCTCGGTATG	71.4
Db	9924	TCCTCTGGTTCATATATGATGATATTTTCTCTGTCAGAGATGACCAAGTCTGGTATG	98.5
OY	7141	AGGGTGTCAACCAACCTCTTCTTGTATCTGAAATCTCTCCACCTGAGAGAAATTTCCAG	72.0
Db	9864	AGGGTGTCAACCAACCTCTTGTGTATCTGAAATCTCTCCACCTGAGAGAAATTTCCAG	98.5
OY	7201	CCCAAGATATAGTATATCATTCGGGTCCACAGACCTGGCTAATATGATGGGGTGTTTGAT	72.6
Db	9804	CCCAAGATATAGTATATCATTCGGGTCCACAGACCTGGCTAATATGATGGGGTGTTTGAT	97.45
OY	7261	CCTAATGTATCCCCATGTGAGCACAAGACTTGTGTGGCAGTATGAGAGAGGTCAAGCTTC	73.2
Db	9744	CCTAATGTATCCCAATGTGAGCACAAGACTTGTGTGGCAGTATGAGAGAGGTCAAGCTTC	96.85
OY	7321	AGAGTCAACAAAGACTGATTTCAACTGGATTTGAGAGACCCCACTTTTATAGTGA	73.8
Db	9684	AGAGTCAACAAAGACTGATTTCAAACTGGATTTGAGAGACCCCACTTTTATAGTGA	96.25
OY	7381	CTTATATCTCTGCGAGTCTCTGATCTCTCTCTTAAATGAGGACATTAATCCCACTGG	74.0
Db	9624	CTTATATCTCTGCGAGTCTCTGATCTCTCTCTTAAATGAGGACATTAATCCCACTGG	95.55
OY	7441	CAGGGTGTGGGAGATCAGAGATCAAAACAGCTGTATCACACTGTGTTCTGTCTTC	75.0
Db	9564	CAGGGTGTGGGAGATCAGAGATCAAAACAGCTGTATCACACTGTGTTCTGTCTTC	95.05
OY	7501	AGGGTCAATCAGACTGGGGTTTCTGAGCATGATTAACAACATCCCAAGCTTGGGATCAAGAA	75.6
Db	9504	AGGGTCAATCAGACTGGGGTTTCTGAGCATGATTAACAACATCCCAAGCTTGGGATCAAGAA	94.45
OY	7561	CTGACACCAATCAACGAGCATGAGAGAGACTCTTGTCTACAGACAGACCTGAGCTTCAAG	76.2
Db	9464	CTGACACCAATCAACGAGCATGAGAGAGACTCTTGTCTACAGAGACCTGAGCTTCAAG	93.85
OY	7621	GAGCTGACGTGCATCTGTTCCCTTGTGCGCTGACAGAAACGGCGTTGTGCTCTGGCTC	76.8
Db	9384	GAGCTGACGTGCATCTGTTCCCTTGTGCGCTGACAGAAACGGCGTTGTGCTCTGGCTC	93.25
OY	7681	CTGGGCTGCGCATGCGCAGAGAAAGCTGTCTCAATCTACATCTCTCAACCTGTGTGCGGCG	77.4
Db	9324	CTGGGCTGCGCATGCGCAGAGAAAGCTGTCTCTCAATCTACATCTCTCAACCTGTGTGCGGCG	92.55
OY	7741	GACTTCCCTCTTCTTTAGCGGCACATTAATATGTTGCGCGTTAGCGCTCTCATATATCCGC	78.0
Db	9284	GACTTCCCTCTCTTTAGCGGCACATTAATATGTTGCGCGTTAGCGCTCTCATATATCCGC	92.05
OY	7801	CATCCCATCTCCAAATATCTCAAGTCCGTATGATGACCTTCCCTAATTATATAGGCTAAGC	78.6
Db	9204	CATCCCATCTCCAAATATCTCAAGTCCGTATGATGACCTTCCCTAATTATATAGGCTAAGC	91.45
OY	7861	ATGCTGAACGCGCATGACGACCGAGGCGTCCCTGTCAATCTGTGGCCCATCTGTATCCAC	79.2
Db	9144	ATGCTGAACGCGCATGACGACCGAGGCGTCCCTGTCAATCTGTGGCCCATCTGTATCCAC	90.85
OY	7921	TGCCGCGCGCCCAAGATATCTGTATCATCGGTCAATGTGTGTCTCTGTCTGTGGGCCCTCTCGTG	79.8
Db	9084	TGCCGCGCGCCCAAGATATCTGTGTCAATCGGTCAATGTGTGTCTCTGTCTGTGGGCCCTCTCGTG	90.25
OY	7981	CTGGGGAATATCTCGGAATGATGATGTTCTGTACTCTCGTTATAGGAGTCTAATTTCTGTT	80.4
Db	9024	CTGGGGAATATCTCGGAATGATGATGTTCTGTACTCTCGTTATAGGAGTCTAATTTCTGTT	89.55
OY	8041	TGTTGTGAAACGTCAGATTTCAATTAACAATGCGGTGGCTGTTTTTTTATATGTGTGTTCTC	81.0
Db	8964	TGTTGTGAAACGTCAGATTTCAATTAACAATGCGGTGGCTGTTTTTTTATATGTGTGTTCTC	89.05
OY	8101	TGTTGGTTCACGCTGTGTCTCTGTGTCAAGATTTCTCTGTGATATCCCGAAGATGCGCGTG	81.6
Db	8904	TGTTGGTTCACGCTGTGTCTCTGTGTCAAGATTTCTCTGTGATATCCCGAAGATGCGCGTG	80.45

QY 8161 ACCGAGCTGACGACCATCTCTCTCAAGTGTCTTCTCTCTCTGAGCTGCC 8220
 DB 8844 ACCAGGCTGACGACCATCTCTCTCAAGTGTCTTCTCTCTCTGAGCTGCC 8785
 QY 8221 TTGGGCAATGAGTGGGCGCTGTTTCCAGATCCACCGATGGAAAGCTTATTTTGT 8280
 DB 8784 TTGGGCAATGAGTGGGCGCTGTTTCCAGATCCACCGATGGAAAGCTTATTTTGT 8725
 QY 8281 CATGTGATCTAGTTTTCATTTTCTCTGCTCTTAAACAGAGTGGCCACCATCATTT 8340
 DB 8724 CATGTGATCTAGTTTTCATTTTCTCTGCTCTTAAACAGAGTGGCCACCATCATTT 8665
 QY 8341 TACTTCTTCTGAGCTCTCTTAAAGCAGGTGAAATAGGCAACCTGAGCTGTTCTC 8400
 DB 8664 TACTTCTTCTGAGCTCTCTTAAAGCAGGTGAAATAGGCAACCTGAGCTGTTCTC 8605
 QY 8401 CAGAGGCTCTGACGACACCGCTGAGGTGATGAGGTGAGGCTGCTCTCAGAA 8460
 DB 8604 CAGAGGCTCTGACGACACCGCTGAGGTGATGAGGTGAGGCTGCTCTCAGAA 8545
 QY 8461 ACCCTGAGCTGTGCGGAGAGATTTGAGAGAGTGAAGACCTCTGCTCTGAGAC 8520
 DB 8544 ACCCTGAGCTGTGCGGAGAGATTTGAGAGAGTGAAGACCTCTGCTCTGAGAC 8485
 QY 8521 AGGACTTTGAGAGAAATGCTGCTGCGCAACCTTGAATATATGATTTTCTTACC 8580
 DB 8484 AGGACTTTGAGAGAAATGCTGCTGCGCAACCTTGAATATATGATTTTCTTACC 8425
 QY 8581 TTTCGCTCAGAAATGCTCTCAGGCTGCC 8609
 DB 8424 TTTCGCTCAGAAATGCTCTCAGGCTGCC 8396

RESULT 3
 AC090099/c 172939 bp DNA linear PRI 01-OCT-2002
 LOCUS Homo sapiens chromosome 11, clone RP11-113D6, complete sequence.
 AC090099
 AC090099.14 GI:23396332
 HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 172939)
 AUTHORS Birren,B., Nusbaum,C. and Lander,E.
 JOURNAL Homo sapiens chromosome 11, clone RP11-113D6
 REFERENCE 2 (bases 1 to 172939)
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
 Barta,N., Bastien,V., Boguslavsky,L., Boukhalter,B., Brown,A.,
 Camarata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Fardo,S., Fero,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
 Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heatford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Karatas,A., Laroque,K., Lamaszates,R., Landers,T.,
 Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
 Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,
 McPheters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
 Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
 Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
 Sougnez,C., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A.,
 Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.

TITLE
 JOURNAL Submitted (14-FEB-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE 3 (bases 1 to 172939)
 AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barta,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Fardo,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gind,S., Graham,L., Grand-Pierre,N., Hagos,B.,
 Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
 Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Talamas,J.,
 Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.

TITLE
 JOURNAL Direct Submission
 REFERENCE Submitted (05-SEP-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 172939)
 AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barta,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Fardo,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gind,S., Graham,L., Grand-Pierre,N., Hagos,B.,
 Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,
 Smith,C., Spencer,B., Strange-Thomann,N., Stojanovic,N., Talamas,J.,
 Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.

TITLE
 JOURNAL Direct Submission
 COMMENT Submitted (01-OCT-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Oct 1, 2002 this sequence version replaced gi:22726159.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 Direct Submission
 Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIRB
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: L12764
 Center clone name: L13_D_6
 FEATURES
 source
 1. 172939
 /location=Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP11-113D6"
 /clone_1lb="RP11-113D6 Human Male BAC"
 /complement (1..77)
 /rpt_family="Alus"
 /complement (154..680)
 /rpt_family="Char11e1"
 /complement (678..810)
 /rpt_family="Char11e1"
 /complement (879..1057)
 repeat_region
 repeat_region
 repeat_region
 repeat_region

```

repeat_region /rpt_family="L1M4"
                complement(1395..1513)
                /rpt_family="FLAM_A"
unsure        1870..1876
                /note="<30 qual SNGL region"
                1902..1910
                /note="<30 qual SNGL region"
repeat_region 2672..2788
                /rpt_family="5S"
                complement(4009..4170)
repeat_region /rpt_family="MIR"
                complement(4210..4355)
repeat_region /rpt_family="MERSB"
                4356..4647
                /rpt_family="AluX"
                complement(6554..6658)
repeat_region /rpt_family="LTR41"
                6916..6985
                /rpt_family="MLT1C"
                complement(7175..7359)
repeat_region /rpt_family="LTR41"
                /rpt_family="Aluub"
                complement(8308..8611)
repeat_region /rpt_family="Aluub"
                8622..9083
                /rpt_family="MLT2B1"
                complement(9265..10668)
repeat_region /rpt_family="L1PB3"
                10699..10809
                /rpt_family="MERS34"
                /rpt_family="L1264"
                10782..11264
                /rpt_family="LTR48B"
                complement(11344..11683)
repeat_region /rpt_family="L2"
                11835..12026
                /rpt_family="L1MC4a"
                12188..12660
                /rpt_family="L1MA2"
                12879..13232
                /rpt_family="L1MC4a"
                13223..13522
                /rpt_family="AluX"
                13523..13817
                /rpt_family="L1MC4a"
                14819..15108
                /rpt_family="AluO"
                15114..15175
                /rpt_family="GAAA)n"
                16150..16336
                /rpt_family="MERSA"
                complement(16587..16598)
                /note="<30 qual SNGL region"
                complement(16607..16612)
                /note="<30 qual SNGL region"
                complement(16746..17360)
                /rpt_family="L1ME3"
                17361..17664
                /rpt_family="AluO"
                complement(17665..17774)
                /rpt_family="L1ME3"
                complement(17827..17929)
                /rpt_family="MIR"
                18993..19298
                /rpt_family="AluO"
                19755..19821
                /rpt_family="L2"
                19933..19941
                /note="<30 qual SNGL region"
                20320..20609
                /rpt_family="AluO"
                complement(20684..20994)
                /rpt_family="AluX"
                20713..20721
                /note="<30 qual SNGL region"

```

```

unsure        20879..20883
                /note="<30 qual SNGL region"
                complement(21118..21398)
                /rpt_family="AluX"
unsure        21141..21147
                /note="<30 qual SNGL region"
                21144
                /note="probably A; possibly G"
unsure        21146
                /note="probably A; possibly G"
unsure        21153..21158
                /note="probably A; possibly G"
unsure        21154
                /note="<30 qual SNGL region"
unsure        21154
                /note="probably T; possibly C"
                21160..21167
                /note="<30 qual SNGL region"

Query Match 99.3%; Score 8564.2; DB 9; Length 172939;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 8594; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

QY 1 TGTATGAGCCATGTCACTTTAATRCCAAAACCGAAGAGATTACAAAAAGAAAC 60
    |||||
DB 121497 TGTATGAGCCATGTCACTTTAATRCCAAAACCGAAGAGATTACAAAAAGAAAC 121438

QY 61 TATPAGCAGTACCACTGATGAATPACATGCGAAGATCCCAAAATACCTAGCTAAC 120
    |||||
DB 121437 TATPAGCAGTACCACTGATGAATPACATGCGAAGATCCCAAAATACCTAGCTAAC 121378

QY 121 CCATCCAAACAGCATATCAAGAGATTAATCCACATGTGCAATGGGTTTCATPACAGG 180
    |||||
DB 121377 CCATCCAAACAGCATATCAAGAGATTAATCCACATGTGCAATGGGTTTCATPACAGG 121318

QY 181 GTGCGAGATAGGTTAATCAATACCAAGTCAATTAATGTGATACATCAATTAACAGATTA 240
    |||||
DB 121317 GTGCGAGATAGGTTAATCAATACCAAGTCAATTAATGTGATACATCAATTAACAGATTA 121258

QY 241 AAAACAAAATACATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAAATCTAAC 300
    |||||
DB 121257 AAAACAAAATACATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAAATCTAAC 121198

QY 301 ATTTCTTATGATTAATAACCTTCAGCAAAATGACATAGAAAGACATACCTTAATGTAA 360
    |||||
DB 121197 ATTTCTTATGATTAATAACCTTCAGCAAAATGACATAGAAAGACATACCTTAATGTAA 121138

QY 361 TAAAAGCCATATATGACGAGCCACAGCAAAATTAATTAATGAGGAAAAAGTTGAAA 420
    |||||
DB 121137 TAAAAGCCATATATGACGAGCCACAGCAAAATTAATTAATGAGGAAAAAGTTGAAA 121078

QY 421 CATTCGCCGTGAGAACTGGAACAAGAGATGTAATCTTACACCACTTCTATTCAACAT 480
    |||||
DB 121077 CATTCGCCGTGAGAACTGGAACAAGAGATGTAATCTTACACCACTTCTATTCAACAT 121018

QY 481 AGTAGTGAAGTTTATAGCCAGAGCAATCAGACAAGAAAGAAATCAAGGGCAACCAAT 540
    |||||
DB 121017 AGTAGTGAAGTTTATAGCCAGAGCAATCAGACAAGAAAGAAATCAAGGGCAACCAAT 120958

QY 541 CATATAAGAGAGTAAATCTGTCTGTCATGATGATTAATGATTTGATTTCTTGAAGA 600
    |||||
DB 120957 CATATAAGAGAGTAAATCTGTCTGTCATGATGATTAATGATTTGATTTCTTGAAGA 120898

QY 601 CCTTAAGACTCATCCAGAAAGCTCTGAATCTGATATCAATTAATCAATTAAGTTTCAG 660
    |||||
DB 120897 CCTTAAGACTCATCCAGAAAGCTCTGAATCTGATATCAATTAATCAATTAAGTTTCAG 120838

QY 661 ATACAACTAAATGTACAAATACAGTAGCACTGTATACCAACCAAGTGA 720
    |||||
DB 120837 ATACAACTAAATGTACAAATACAGTAGCACTGTATACCAACCAAGTGA 120778

QY 721 GAATCAATCAAGAACTCAACACTTTTACATAGCTGTAATAAATTAATTAAGATATT 780
    |||||
DB 120777 GAATCAATCAAGAACTCAACACTTTTACATAGCTGTAATAAATTAATTAAGATATT 120718

```


Qy	781	CTTACCAAGAGGTGAAGACCTCTCAAGGAAAACTACAGAAACACAGCTGACATATA	840
Db	12017	CTTACCAAGAGGTGAAGACCTCTCAAGGAAAACTACAGAAACACAGCTGACATATA	120658
Qy	841	GATGACCAAAAGTGGAAACATCCATGCTCATGATGGGTAGATCAATATTGTG	900
Db	12057	GATGACCAAAAGTGGAAACATCCATGCTCATGATGGGTAGATCAATATTGTG	120598
Qy	901	AAATGACCAATTTGGCAAAAGCAATCTACAGTTCAATGCAATTTCCACCAAAATATCA	960
Db	12097	AAATGACCAATTTGGCAAAAGCAATCTACAGTTCAATGCAATTTCCACCAAAATATCA	120538
Qy	961	TCATCATCTTCTCAGAACTAGAAAAACAATCTTAATAATTCCTATGAAACCAACCA	1020
Db	120537	TCATCATCTTCTCAGAACTAGAAAAACAATCTTAATAATTCCTATGAAACCAACCA	120479
Qy	1021	AAAAAAAAAAAAAAAACCCGCATAGCCAAAGCAAGACTTAGCAAAAAGAACTATGGA	1080
Db	120478	AAAAAAAAAAAAAAAACCCGCATAGCCAAAGCAAGACTTAGCAAAAAGAACTATGGA	120419
Qy	1081	GGCATCACTTACCCATCTTCAAACTATACAGGCTATATACCAAAACATCATGG	1140
Db	120418	GGCATCACTTACCCATCTTCAAACTATACAGGCTATATACCAAAACATCATGG	120359
Qy	1141	CATGACATATAAAGTAGGACATPAGCAATGGAAGAAAGAGAAATCCAGAAATTAAG	1200
Db	120358	CATGACATATAAAGTAGGACATPAGCAATGGAAGAAAGAGAAATCCAGAAATTAAG	120299
Qy	1201	CCAAATTAATATAGCCCAATGATTTTGAACAAGCAAAACAAATTAAGTGGGAAA	1260
Db	120298	CCAAATTAATATAGCCCAATGATTTTGAACAAGCAAAACAAATTAAGTGGGAAA	120239
Qy	1261	GACATCTAGTTAACAAATGCTGTGATTAATTTGCAAGCCCATGTGGAAGATGAAA	1320
Db	120238	GACATCTAGTTAACAAATGCTGTGATTAATTTGCAAGCCCATGTGGAAGATGAAA	120179
Qy	1321	CTGATCTCCCTGTCTCTCACTTAATACAAAAATTGACACAAAGATGATCAAAAGCTTAA	1380
Db	120178	CTGATCTCCCTGTCTCTCACTTAATACAAAAATTGACACAAAGATGATCAAAAGCTTAA	120120
Qy	1381	TCGAGACCTAAACCAATAAAATTTCTAGAAATACATGAGAAATGCTTCTAGACAT	1440
Db	120119	TCGAGACCTAAACCAATAAAATTTCTAGAAATACATGAGAAATGCTTCTAGACAT	120060
Qy	1441	TCATTAAGCAAAAGCTTCAATGCGCAAGAACCCAAAGTAAATGCAACAAAAAAT	1500
Db	120059	TCATTAAGCAAAAGCTTCAATGCGCAAGAACCCAAAGTAAATGCAACAAAAAAT	120000
Qy	1501	AAATAGATAGACTTAATTTAACTTAAAGCTTTTGGCAGCAAAAAACATCATAGCAG	1560
Db	119999	AAATAGATAGACTTAATTTAACTTAAAGCTTTTGGCAGCAAAAAACATCATAGCAG	119940
Qy	1561	AGCAAAACAGAACCCCAAGAGTAGAGAAATCTTCAAAACATAGCATCTGACTAGG	1620
Db	119993	AGCAAAACAGAACCCCAAGAGTAGAGAAATCTTCAAAACATAGCATCTGACTAGG	119880
Qy	1621	ACTAATATCGGAATCCACAAGAACTCAAAACAATCAGCAGAGAGAAAGCAAACTCC	1680
Db	119879	ACTAATATCGGAATCCACAAGAACTCAAAACAATCAGCAGAGAGAAAGCAAACTCC	119820
Qy	1681	CATGAAAGAGTGGGCTPAGGACATGAAATAGCAATTTCTCAAAAGAAATTAACAATGGC	1740
Db	119819	CATGAAAGAGTGGGCTPAGGACATGAAATAGCAATTTCTCAAAAGAAATTAACAATGGC	119760
Qy	1741	CAACAAACAGAAAAAATGCTTAACATCACTAATGTTAGGAAATGTAATCAACCTG	1800
Db	119759	CAACAAACAGAAAAAATGCTTAACATCACTAATGTTAGGAAATGTAATCAACCTG	119700
Qy	1801	TAAATGCAATCACTTATCTCTGCAAGAAATGTCATTAATTTTAAATTTTAAATATA	1860
Db	119699	TAAATGCAATCACTTATCTCTGCAAGAAATGTCATTAATTTTAAATTTTAAATATA	119640
Qy	1861	GATGTTGGTGGTCTGTGTGTATAAAGAACATTTTACACTGTGTGGGAATGTAAAC	1920

Db	119639	GATGTTGGTGGTCTGTGTGTATAAAGAACATTTTACACTGTGTGGGAATGTAAAC	119580
Qy	1921	TTGGCAACCACTATGAAAAAGTGTGAATTTCTTAAAGAACTAAAGTAGATGAC	1980
Db	119579	TTGGCAACCACTATGAAAAAGTGTGAATTTCTTAAAGAACTAAAGTAGATGAC	119520
Qy	1981	CATTGATCCAGCAATCCATTAATATGTATTAATATATATATTTATATCCATGGAAT	2040
Db	119519	CATTGATCCAGCAATCCATTAATATGTATTAATATATATATTTATATCCATGGAAT	119460
Qy	2041	ACAACCTGAGCAATAAAAAATTAATGATGACATTCACAGCAATCTAGTGAATTTG	2100
Db	119459	ACAACCTGAGCAATAAAAAATTAATGATGACATTCACAGCAATCTAGTGAATTTG	119400
Qy	2101	AGACCTTATCTTAAGTGGGTAACCTCAGAAATGAAAAACCAACATCATATGTTCTAC	2160
Db	119399	AGACCTTATCTTAAGTGGGTAACCTCAGAAATGAAAAACCAACATCATATGTTCTAC	119340
Qy	2161	TTACAAATGGGGCTAAGCTGTGAGACACGAAAGCATPAGATGATTAATGAATCTGG	2220
Db	119339	TTACAAATGGGGCTAAGCTGTGAGACACGAAAGCATPAGATGATTAATGAATCTGG	119280
Qy	2221	GGACTTGAAGGGAAGGATGGAAGAGAGGAGGTAAGAGACTACCAATGGGTACAGT	2280
Db	119279	GGACTTGAAGGGAAGGATGGAAGAGAGGAGGTAAGAGACTACCAATGGGTACAGT	119220
Qy	2281	GTCACCTGCTCAGGTGATGGGTGACCAAAATCTCAGAAATTAACCTAAAGAACTTATC	2340
Db	119219	GTCACCTGCTCAGGTGATGGGTGACCAAAATCTCAGAAATTAACCTAAAGAACTTATC	119160
Qy	2341	CATGAAAGCAACCACTGTTCCCAAAATCCCAATGAAATTAATAATTAATTA	2400
Db	119159	CATGAAAGCAACCACTGTTCCCAAAATCCCAATGAAATTAATAATTAATTAATTA	119100
Qy	2401	ATGATTTAATTTTCAACAATTTAAAAAAGTTCAGTCTGATGAGTTTAAATTAAGATTA	2460
Db	119099	ATGATTTAATTTTCAACAATTTAAAAAAGTTCAGTCTGATGAGTTTAAATTAAGATTA	119040
Qy	2461	AGATGAAAAAGTAGCAAGTGTAGCTCTGACATGGAACATGGAACCTAGATTTTACCTTG	2520
Db	119039	AGATGAAAAAGTAGCAAGTGTAGCTCTGACATGGAACATGGAACCTAGATTTTACCTTG	118980
Qy	2521	CATACCTTCTCTGCAATTTGAAAGAAATATACATTAATATATACCATATATATA	2580
Db	118979	CATACCTTCTCTGCAATTTGAAAGAAATATACATTAATATATACCATATATATA	118920
Qy	2581	AAAGAAACATCAAAAAATTTGTCAAGGCCATTTGACCTTGAATGCTCCATGATATAC	2640
Db	118919	AAAGAAACATCAAAAAATTTGTCAAGGCCATTTGACCTTGAATGCTCCATGATATAC	118860
Qy	2641	TTTTTCAATTTGGAATATTAAGCTCATATAGATGATGATGATGATGATGATGATG	2700
Db	118859	TTTTTCAATTTGGAATATTAAGCTCATATAGATGATGATGATGATGATGATGATG	118800
Qy	2701	GCCCATTAATTTTAAAAATCAGACAGCAATTAATTAACAGAACTTGAACAAAGTGG	2760
Db	118799	GCCCATTAATTTTAAAAATCAGACAGCAATTAATTAACAGAACTTGAACAAAGTGG	118740
Qy	2761	GACGTGATAGAGAGGCTTAAATATGTATGGAATGGAACAGAGCCAAAGATCTCTCAG	2820
Db	118739	GACGTGATAGAGAGGCTTAAATATGTATGGAATGGAACAGAGCCAAAGATCTCTCAG	118680
Qy	2821	TTAGATTTGTCTCAAAATACCTGAGCTCTGATTTGGCCATAGCTCATACAGAA	2880
Db	118679	TTAGATTTGTCTCAAAATACCTGAGCTCTGATTTGGCCATAGCTCATACAGAA	118620
Qy	2881	ATAACAAAGCTGTCCAGATCTTGTATAGCTTGATGATTTCACTTCAATTTCACT	2940
Db	118619	ATAACAAAGCTGTCCAGATCTTGTATAGCTTGATGATTTCACTTCAATTTCACT	118560
Qy	2941	CTGTAGGCACTCTCGAATTTAAGCAACAGAAAGTCTCTGAACTCACTGAATCCCA	3000

Db 118559 CCTGTAGGATCTCTGTAATTAGACAACAGAAAAGTCTCTGAGATGACTGAATCCCA 118500
 Qy 3001 GAAAGGCTCTCTACTTTAGACAAGAGGCTCTTCAACACTGAGCAAAAGAGAAAGAT 3060
 Db 118499 TAAAGGCTCTCTACTTTAGACAAGAGGAGCTCTTCAACACTGAGCAAAAGAGAAAGAT 118440
 Qy 3061 AAGGTAAGTACCAAGAACTCTCTTCAAGAGTATGATTTTGGCTGTAAGATC 3120
 Db 118439 AAGGTAAGTACCAAGAACTCTCTTCAAGAGTATGATTTTGGCTGTAAGATC 118380
 Qy 3121 ATGCTTATGCTTCACTTCTGCTGCTACATGAGGGGGTCAAGAGCTTGTTCAGAAA 3180
 Db 118379 ATGCTTATGCTTCACTTCTGCTGCTACATGAGGGGGTCAAGAGCTTGTTCAGAAA 118320
 Qy 3181 AGACAGAGACATGAAAGCTTCTTTCAGAACTGAGTCTGTCAACCCAACTGTGTGAG 3240
 Db 118319 AGACAGAGACATGAAAGCTTCTTTCAGAACTGAGTCTGTCAACCCAACTGTGTGAG 118260
 Qy 3241 CTCTAAATGAGTGTCCCTCTTAAATTTATCTCCCATATACCTGCTTCAATTCAGATCA 3300
 Db 118259 CTCTAAATGAGTGTCCCTCTTAAATTTATCTCCCATATACCTGCTTCAATTCAGATCA 118200
 Qy 3301 TTCAATCTGCTCTCACTGAGAGAGCTGCTCTTCAATTCATTTAAGAGAGAGGAG 3360
 Db 118199 TTCAATCTGCTCTCACTGAGAGAGCTGCTCTTCAATTCATTTAAGAGAGAGGAG 118140
 Qy 3361 ATGCAAGCATTTCTTCCAGAGTGAACGTCTATAGAGCAAGTTTCTTGTTCACCTAC 3420
 Db 118139 ATGCAAGCATTTCTTCCAGAGTGAACGTCTATAGAGCAAGTTTCTTGTTCACCTAC 118080
 Qy 3421 TTTTCAATTTATCTCTTGTGCTATCTGAAAAGGCTTAAGAGATATAGATGAGCCCA 3480
 Db 118079 TTTTCAATTTATCTCTTGTGCTATCTGAAAAGGCTTAAGAGATATAGATGAGCCCA 118020
 Qy 3481 ATAATTTAAGAGTGTTCATGAGAAAGTATTTACAAAGTGCACAGAGTTAAGAGTCA 3540
 Db 118019 ATAATTTAAGAGTGTTCATGAGAAAGTATTTACAAAGTGCACAGAGTTAAGAGTCA 117960
 Qy 3541 GATCTTAAGAGCATATCATATAGGGAGCATCTTCTCCCTTAAGGCTGAAGAGAGAG 3600
 Db 117959 GATCTTAAGAGCATATCATATAGGGAGCATCTTCTCCCTTAAGGCTGAAGAGAGAG 117900
 Qy 3601 GAAGAGAGCATATCATATAGGGAGCATCTTCTCCCTTAAGGCTGAAGAGAGAGAT 3660
 Db 117899 GAAGAGAGCATATCATATAGGGAGCATCTTCTCCCTTAAGGCTGAAGAGAGAT 117840
 Qy 3661 GAGCAGGCAAGTGAAGAAAGCCCTGCGTGCCAAAGCAGCAGCAGAGCTGATATAGT 3720
 Db 117839 GAGCAGGCAAGTGAAGAAAGCCCTGCGTGCCAAAGCAGCAGCAGAGCTGATATAGT 117780
 Qy 3721 TGGATCTGTCTCCCAACCAAAATCTCATGTGATTTGTAATTTCCAAATGTTGAGAAAG 3780
 Db 117779 TGGATCTGTCTCCCAACCAAAATCTCATGTGATTTGTAATTTCCAAATGTTGAGAAAG 117720
 Qy 3781 CCTTGGAGAGATGATTTAGATGACGAGGAGATGTTTGGATGAAATGTTTAAACCAT 3840
 Db 117719 CCTTGGAGAGATGATTTAGATGACGAGGAGATGTTTGGATGAAATGTTTAAACCAT 117660
 Qy 3841 CCCCTTGGATGATTTGTTGTTGATGATGACGAGTTCTCATGAAATCTAGTTGTTTAA 3900
 Db 117659 CCCCTTGGATGATTTGTTGTTGATGATGACGAGTTCTCATGAAATCTAGTTGTTTAA 117600
 Qy 3901 GGGTGTAGACCTCTCTCTCTCTCTTACTCTGCTCTCAAGATGAGAGCGCTGCTC 3960
 Db 117599 GGGTGTAGACCTCTCTCTCTCTCTCTTACTCTGCTCTCAAGATGAGAGCGCTGCTC 117540
 Qy 3961 CCCCTTGGCTTTACACAGAGATGAAAGCTTCTGAGAGCTCCCAAGAGAGAGAGCTGC 4020
 Db 117539 CCCCTTGGCTTTACACAGAGATGAAAGCTTCTGAGAGCTCCCAAGAGAGAGAGCTGC 117480
 Qy 4021 TATGCTTTCTGTATGACTGTGTAGAGCTATTAAGCAGTTAAACCATTTCTTCAATAAT 4080
 Db 117479 TATGCTTTCTGTATGACTGTGTAGAGCTATTAAGCAGTTAAACCATTTCTTCAATAAT 117420

Qy 4081 TCCAGCTCAGGATATTTCTTTTAGCAATTTGAGAAATGAATCTATATACAGAGAGAG 4140
 Db 117419 TCCAGCTCAGGATATTTCTTTTAGCAATTTGAGAAATGAATCTATATACAGAGAGAG 117360
 Qy 4141 CCAGAGATGAGAAATCCCAAGGCTTCTCTGCTCTCTCCAGTCTCTGCTGTGTCTC 4200
 Db 117359 CCAGAGATGAGAAATCCCAAGGCTTCTCTGCTCTCTCTCCAGTCTCTGCTGTGTCTC 117300
 Qy 4201 CCAGTGTCTCAATTCACAGAAACAGAAATTAAGATCCCATGATGTGTACTTA 4260
 Db 117299 CCAGTGTCTCAATTCACAGAAACAGAAATTAAGATCCCATGATGTGTACTTA 117240
 Qy 4261 GAAAGCATCTCTGGAGATGTAAGCAGGATTAAGAGATGAGAAAGAAATCCCATAG 4320
 Db 117239 GAAAGCATCTCTGGAGATGTAAGCAGGATTAAGAGATGAGAAAGAAATCCCATAG 117180
 Qy 4321 TAAATGAGACTATCCCTCTCACTTCTGATCTCTTAATTCCTGAGGCTTCTCTATC 4380
 Db 117179 TAAATGAGACTATCCCTCTCACTTCTGATCTCTTAATTCCTGAGGCTTCTCTATC 117120
 Qy 4381 TGAATGATCCCTGTCTCAATTCAGCTCTATCAGACTACTTAAATGTTGGCTTCTTTC 4440
 Db 117119 TGAATGATCCCTGTCTCAATTCAGCTCTATCAGACTACTTAAATGTTGGCTTCTTTC 117060
 Qy 4441 TCTACTGCACTTTATGAGAAATGTTGCAATTTGTTAAATGCAATGAAATTAAT 4500
 Db 117059 TCTACTGCACTTTATGAGAAATGTTGCAATTTGTTAAATGCAATGAAATTAAT 117000
 Qy 4501 GTATATTTAAAGAAACATATGATTTTGTATGATATAGTTTGGCTGATCTAATAA 4560
 Db 116999 GTATATTTAAAGAAACATATGATTTTGTATGATATAGTTTGGCTGATCTAATAA 116940
 Qy 4561 GACATGAGAGAAATATCTTAAACAGAAAGTATGTTGCTCTGCTGATCTAGATT 4620
 Db 116939 GACATGAGAGAAATATCTTAAACAGAAAGTATGTTGCTCTGCTGATCTAGATT 116880
 Qy 4621 CTGAATCTCAGATTCAACAACTACAGAGAGAACTTTCCAAATTAAGAGTGTGGC 4680
 Db 116879 CTGAATCTCAGATTCAACAACTACAGAGAGAACTTTCCAAATTAAGAGTGTGGC 116820
 Qy 4681 GAGTGTGATGATCTGAACAGGTAACAACTGTATTTCTTGTGATTAATTTCTGAAAA 4740
 Db 116819 GAGTGTGATGATCTGAACAGGTAACAACTGTATTTCTTGTGATTAATTTCTGAAAA 116760
 Qy 4741 CTACATATTAACAGAACTTATATAGCAATTTGCAATTTGTCAATTCTAATACTTT 4800
 Db 116759 CTACATATTAACAGAACTTATATAGCAATTTGCAATTTGTCAATTCTAATACTTT 116700
 Qy 4801 AATGATTTAATGATCTGAGAGAAAGTGCATAGATATACAAATACCATATTAAG 4860
 Db 116699 AATGATTTAATGATCTGAGAGAAAGTGCATAGATATACAAATACCATATTAAG 116640
 Qy 4861 AATGATGATCTGAGATTTTGTCTGTCTGAGGCTTCTGAGAAATCCCTGTAAT 4920
 Db 116639 AATGATGATCTGAGATTTTGTCTGTCTGAGGCTTCTGAGAAATCCCTGTAAT 116580
 Qy 4921 ACACAAAAATGACACTTCTGAGATCTGAATCTGAAGCTCCAAAGCATATCATACAGA 4980
 Db 116579 ACACAAAAATGACACTTCTGAGATCTGAATCTGAAGCTCCAAAGCATATCATACAGA 116520
 Qy 4981 TTCCAAAAATGCTGCTCCCAAGTCTCTGAGAGTGGCTCATGCTTGTGATCTTACAT 5040
 Db 116519 TTCCAAAAATGCTGCTCCCAAGTCTCTGAGAGTGGCTCATGCTTGTGATCTTACAT 116460
 Qy 5041 GGTTCACAGGACATTAGCATTCAGTCTTATGAGAAAAAGAGAGGAGAGAGGCT 5100
 Db 116459 GGTTCACAGGACATTAGCATTCAGTCTTATGAGAAAAAGAGAGGAGAGAGGCT 116400
 Qy 5101 TTGCTCTTCTATTAATCCCATGAGCAGAGCTTGTCTGTCACTTTTGTGATTTCTTC 5160
 Db 116399 TTGCTCTTCTATTAATCCCATGAGCAGAGCTTGTCTGTCACTTTTGTGATTTCTTC 116340


```
QY 5161 ACTTAACGACCTGCTCATGGGATGTCATCCAGATTCAGAAACTGGGATGCGGTC 5220
Db 11639 ACTTAACGACCTGCTCATGGGATGTCATCCAGATTCAGAAACTGGGATGCGGTC 116280
QY 5221 CTGTGCTGCTTGATACATTCAGAAAGTATGAGCAAAAAGAAATCTGGGGCA 5280
Db 116279 CTGTGCTGCTTGATACATTCAGAAAGTATGAGCAAAAAGAAATCTGGGGCA 116220
QY 5281 ACAGCAGTCTCTTCAAGCCCTGACCTGCTCTGATCTGTGCTCAATCAAAATTTTCA 5340
Db 116219 ACAGCAGTCTCTTCAAGCCCTGACCTGCTCTGATCTGTGCTCAATCAAAATTTTCA 116160
QY 5341 GGAATCTCTAGAAATTAATATGTTGGGAGAGAAAGAACTGGAGTCTCGTCAGGA 5400
Db 116159 GGAATCTCTAGAAATTAATATGTTGGGAGAGAAAGAACTGGAGTCTCGTCAGGA 116100
QY 5401 CTCAGGAGACCAAGGGGCTGTATGAGACCTGCTCTTCAATGTTGAAACAGAGAAACCT 5460
Db 116099 CTCAGGAGACCAAGGGGCTGTATGAGACCTGCTCTTCAATGTTGAAACAGAGAAACCT 116040
QY 5461 TTAATCTCTAGGCTTACCTGATCTTATGTTATGAGAAATACATAGACAGTCT 5520
Db 116039 TTAATCTCTAGGCTTACCTGATCTTATGTTATGAGAAATACATAGACAGTCT 115980
QY 5521 TTAAGAACTCAATGATGATGTTAAACAAATGCTTAAATGTTGATACAGAGTAA 5580
Db 115979 TTAAGAACTCAATGATGATGTTAAACAAATGCTTAAATGTTGATACAGAGTAA 115920
QY 5581 AGACAGCATGATCTTACTCTTACGAGCTTCCGGGTTTATGAGAGAAACAAATATCA 5640
Db 115919 AGACAGCATGATCTTACTCTTACGAGCTTCCGGGTTTATGAGAGAAACAAATATCA 115860
QY 5641 TACCAATCTATAGATGAGCAAAAGTTAGTGTCTGAGTGTGATTAACAGAGTCTC 5700
Db 115859 TACCAATCTATAGATGAGCAAAAGTTAGTGTCTGAGTGTGATTAACAGAGTCTC 115800
QY 5701 CTTTTCTCCCATTTCTTTTGGGCAATCAGAGCTGTGAGCTGTCTTCTTAAGAG 5760
Db 115799 CTTTTCTCCCATTTCTTTTGGGCAATCAGAGCTGTGAGCTGTCTTCTTAAGAG 115740
QY 5761 AGCTCATGATGATGATGATCTCTGATGCTCTCTATATCTCCAGAGAGATGATC 5820
Db 115739 AGCTCATGATGATGATGATCTCTGATGCTCTCTATATCTCCAGAGAGATGATC 115680
QY 5821 TTTCTTCACTGAGAGCTCTGCGCAATGTCATTTCTGGGATTCAGAGCAACGTG 5880
Db 115679 TTTCTTCACTGAGAGCTCTGCGCAATGTCATTTCTGGGATTCAGAGCAACGTG 115620
QY 5881 CCTCTGATAGGCAAAAGAACTCTGAAATTTGTTCTTAATGGCAGCATCACTCTA 5940
Db 115619 CCTCTGATAGGCAAAAGAACTCTGAAATTTGTTCTTAATGGCAGCATCACTCTA 115560
QY 5941 TTTTTCCCTATTTTCAATTTGCTTCAATCTCTATCTGAGATTTGTTAGGTTAATTTT 6000
Db 115559 TTTTTCCCTATTTTCAATTTGCTTCAATCTCTATCTGAGATTTGTTAGGTTAATTTT 115500
QY 6001 TTTTTCAGCCCAATTTTGAAGTCAATCTGAAATTTAACTTGAAGATCACTCTCTA 6060
Db 115499 TTTTTCAGCCCAATTTTGAAGTCAATCTGAAATTTAACTTGAAGATCACTCTCTA 115440
QY 6061 TTAACCCCTCTATCATGTAATATGACACATATGTTGCTGGGTCAGAAAGGCTGTGA 6120
Db 115439 TTAACCCCTCTATCATGTAATATGACACATATGTTGCTGGGTCAGAAAGGCTGTGA 115380
QY 6121 AAAAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6180
Db 115379 AAAAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 115320
QY 6181 CCGGAGCAAAATATATCTCTCTTATATCTTAAATTTTCAATATCTGCAATACAGAAAT 6240
Db 115319 CCGGAGCAAAATATATCTCTCTTATATCTTAAATTTTCAATATCTGCAATACAGAAAT 115260
QY 6241 TACTAATTTTAACTCTAGGCTCTTCAATGATTTAAAGGCAATACCTAATTAATG 6300
Db 115259 TACTAATTTTAACTCTAGGCTCTTCAATGATTTAAAGGCAATACCTAATTAATG 115200
QY 6301 TCAATCAGCTGCTGTATATCTCCCAATTAGACCTAATCTCATTTCTCAGTTGAAAT 6360
Db 115199 TCAATCAGCTGCTGTATATCTCCCAATTAGACCTAATCTCATTTCTCAGTTGAAAT 115140
QY 6361 GCATGAATATCTCTTTTAAACCCAGCCCTTACATTTCTCTATTTTCACTCATGAGCT 6420
Db 115139 GCATGAATATCTCTTTTAAACCCAGCCCTTACATTTCTCTATTTTCACTCATGAGCT 115080
QY 6421 CCTCTCATGAAAGTTTGAATGCAACAAAGAGCTAACAAATATCTCCGAAAGAG 6480
Db 115079 CCTCTCATGAAAGTTTGAATGCAACAAAGAGCTAACAAATATCTCCGAAAGAG 115020
QY 6481 AATGAATAGGTTTACATTTGATGATGATGATGATGATGATGATGATGATGATGAT 6540
Db 115019 AATGAATAGGTTTACATTTGATGATGATGATGATGATGATGATGATGATGATGAT 114960
QY 6541 TCCCACTTCAATTTACCTGCTGAGTGGCACTGAGCTCACCTCACTTACTCTTCT 6600
Db 114959 TCCCACTTCAATTTACCTGCTGAGTGGCACTGAGCTCACCTCACTTACTCTTCT 114900
QY 6601 CTGTTCTATGCTGAGGAAATTAAGCTGAGACCCACACCTTACCCAACTGTGTACAA 6660
Db 114899 CTGTTCTATGCTGAGGAAATTAAGCTGAGACCCACACCTTACCCAACTGTGTACAA 114840
QY 6661 ATGCTTACAGGGGTTTCGGCAAGCACTGAGTCTTATTTTAAAGGCATGATGCT 6720
Db 114839 ATGCTTACAGGGGTTTCGGCAAGCACTGAGTCTTATTTTAAAGGCATGATGCT 114780
QY 6721 CAATTTCAAGTTTGGGCACTCATCAATCTTCTTCAACACAGATGAGCTGTCCA 6780
Db 114779 CAATTTCAAGTTTGGGCACTCATCAATCTTCTTCAACACAGATGAGCTGTCCA 114720
QY 6781 AATGAATCTGAGAAATGAAATTTTCTCATCTAATTAATGATGATGATGATGATGAT 6840
Db 114719 AATGAATCTGAGAAATGAAATTTTCTCATCTAATTAATGATGATGATGATGATGAT 114660
QY 6841 TACATTTGCTTCAATTTTATTTTCAATTTCAATTTCAATTTCAATTTCAATTTCAAT 6900
Db 114659 TACATTTGCTTCAATTTTATTTTCAATTTCAATTTCAATTTCAATTTCAATTTCAAT 114600
QY 6901 CTCATGATGATCTTCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 6960
Db 114599 CTCATGATGATCTTCAATTTGATGATGATGATGATGATGATGATGATGATGATGAT 114540
QY 6961 GAATGAGCATGAGAGATTAATTAATGATGATGATGATGATGATGATGATGATGATGAT 7020
Db 114539 GAATGAGCATGAGAGATTAATTAATGATGATGATGATGATGATGATGATGATGATGAT 114480
QY 7021 GAGTTATTTTCTACCCACAGGTCCTGCTTCTGATGATGATGATGATGATGATGATGAT 7080
Db 114479 GAGTTATTTTCTACCCACAGGTCCTGCTTCTGATGATGATGATGATGATGATGATGAT 114420
QY 7081 TCTCTGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7140
Db 114419 TCTCTGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 114360
QY 7141 AGGTTTCAACCACTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 7200
Db 114359 AGGTTTCAACCACTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 114300
QY 7201 CCGAGATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7260
Db 114299 CCGAGATGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 114240
QY 7261 CTTAATGTTATCCCATGTCAGACAGAACTTGTGTGAGAGAGAGGTCAGGCTTC 7320
Db 114239 CTTAATGTTATCCCATGTCAGACAGAACTTGTGTGAGAGAGAGGTCAGGCTTC 114180
QY 7321 AGAGTCAACAAAGCTGATTTTCAACTGATTTGAGAGACCCCACTTTTGAATGATG 7380
```

Db	114179	AGAGTCAACAAGAACTGGATTTCAAATCTGGAATTGTAGGACC	CCCACTTTTGATAGTGA	11412
Qy	7381	CTTATTCTTGTGGAGTCTCTGATCTCTCTCTTAAATAGGACA	CACTAAATCCCAATGG	7440
Db	114119	CTTATTCTTGTGGAGTCTCTGATCTCTCTCTTAAATAGGACA	CACTAAATCCCAATGG	114060
Qy	7441	CAGGGTGGGGGAGAAATCAGAGATCAAAACAGCTGGTGAAT	CAATCTGGTTCTGTCTC	7500
Db	114059	CAGGGTGGGGGAGAAATCAGAGATCAAAACAGCTGGTGAAT	CAATCTGGTTCTGTCTC	114000
Qy	7501	AGGGTCAATAGACTGGGGTTTCTGAGACATGATTTCAACCA	TCCAGTCTTGGGTACAAA	7560
Db	113999	AGGGTCAATAGACTGGGGTTTCTGAGACATGATTTCAACCA	TCCAGTCTTGGGTACAAA	113940
Qy	7561	CTGACACCAATCAACGGAGCTGGAGAGACTCTTGGCTAAC	AGAGACCCGAGCTTCAAG	7620
Db	113939	CTGACACCAATCAACGGAGCTGGAGAGACTCTTGGCTAAC	AGAGACCCGAGCTTCAAG	113880
Qy	7621	GGGCTGACGTGATCGTTTCCCTTGTGCGCTGACAGAAA	CGCGGTGTGTCTTGGCTC	7680
Db	113879	GGGCTGACGTGATCGTTTCCCTTGTGCGCTGACAGAAA	CGCGGTGTGTGTCTTGGCTC	113820
Qy	7681	CTGGGCTGGCGGATGGCGAGAAACGCTGTCTCACTCAAT	CTCTCAACCTGTGGCGGCC	7740
Db	113819	CTGGGCTGGCGGATGGCGAGAAACGCTGTCTCACTCAAT	CTCTCAACCTGTGGCGGCC	113760
Qy	7741	GACTTCTTCTTCTTTAGCGGACCAATTATATGTTTGGC	CGTTAGCCTCAATATCCG	7800
Db	113759	GACTTCTTCTTCTTTAGGGGCACTATATATGTTGGCC	GTATAGCCTCAATATCCG	113700
Qy	7801	CATCCCATCTCCAAAATCTCTAGTCTGTGATGATCACT	TTTCCCTTATATAGGCTTAAGC	7860
Db	113699	CATCCCATCTCCAAAATCTCTAGTCTGTGATGATCACT	TTTCCCTTATATAGGCTTAAGC	113640
Qy	7861	ATGCTGAGGCGCATCGACACCGAGCGCTGCTCCATCC	CTGGAGCCATCTGGTACAC	7920
Db	113639	ATGCTGAGGCGCATCGACACCGAGCGCTGCTCCATCC	CTGGAGCCATCTGGTACAC	113580
Qy	7921	TGCGCGCCGCCAGATACCTGTATCATGATCATGTGTG	TCTGTGCGGCCCTGTCCCTG	7980
Db	113579	TGCGCGCCGCCAGATACCTGTATCATGATCATGTGTG	TCTGTGCGGCCCTGTCCCTG	113520
Qy	7981	CTGGGGAAGATCTGGAGTGGAAATGTTCTGTGACTT	CTGTTTAGTGTGATATCTGTT	8040
Db	113519	CTGGGGAAGATCTGGAGTGGAAATGTTCTGTGACTT	CTGTTTAGTGTGATATCTGTT	113460
Qy	8041	TGTTGTGAAACGTCAGATTTCAATCAATCGCTGGCT	GTGTTTTTTATGTGTGTTCTC	8100
Db	113459	TGTTGTGAAACGTCAGATTTCAATCAATCGCTGGCT	GTGTTTTTTATGTGTGTTCTC	113400
Qy	8101	TGTGGGTCCAGCTGGTCTGTCTGTGCAGAGATTCT	CTGTGTGAAATCCCGAAAGTCCGCTG	8160
Db	113399	TGTGGGTCCAGCTGGTCTGTCTGTGCAGAGATTCT	CTGTGTGAAATCCCGAAAGTCCGCTG	113340
Qy	8161	ACCAGGCTGTAGCTGACCAATCCCTCTACAGTGTG	CTTCTCTGTGTGCTGCC	8220
Db	113339	ACCAGGCTGTAGCTGACCAATCCCTCTACAGTGTG	CTTCTCTGTGTGCTGCC	113280
Qy	8221	TTTGGCATTCAGTGGGCGCTGTTTTCAGAGATCCAC	CTTGATTTGAAAGTCTTATTTTGT	8280
Db	113279	TTTGGCATTCAGTGGGCGCTGTTTTCAGAGATCCAC	CTTGATTTGAAAGTCTTATTTTGT	113220
Qy	8281	CATGTGCATCTAGTTCCTATTTTCTGTGCGCTTTAA	CAGAGTGCACAAACCCATCATT	8340
Db	113219	CATGTGCATCTAGTTCCTATTTTCTGTGCGCTTTAA	CAGAGTGCACAAACCCATCATT	113160
Qy	8341	TACTTCTTGTGGGCTCTCTTATAGGACGCTCAAA	ATATAGCGAAGCTGAACTGTCTC	8400
Db	113159	TACTTCTTGTGGGCTCTCTTATAGGACGCTCAAA	ATATATAGCGAAGCTGTCTC	113100
Qy	8401	CAGAGGGCTCTGACAGACACGCTTAGAGTGAATGAA	AGTGTGAAGGTGTCTTCTTCAGAA	8460
Db	113099	CAGAGGGCTCTGACAGACACGCTTAGAGTGAATGAA	AGTGTGAAGGTGTGTCTTCTTCAGAA	113040

QY	8461	ACCGTAGAGCTGTCGGGAGACAGATTGAGAGAGTAGAGAAACACTTCGCGCTGTAGAC	8520
Db	113039	ACCGTAGAGCTGTCGGGAGACAGATTGAGAGAGTAGAGAAACACTTCGCGCTGTAGAC	1129880
QY	8521	AGGACTTGTAGAGACAAATGCTGCGCTGCACCCCTTGACATTAATATGCAATTTTCTTAGCC	8580
Db	112979	AGGACTTGTAGAGACAAATGCTGCGCTGCACCCCTTGACATTAATATGCAATTTTCTTAGCC	1129200
QY	8581	TTCTGCGCTCAGAAATGCTTCAGGGTCCCC	8609
Db	112919	TTCTGCGCTCAGAAATGCTTCAGGTGCTCC	112891
RESULT 4	AC023078/c		
LOCUS	AC023078	163718 bp	DNA linear PRI 23-APR-2002
DEFINITION	Homo sapiens chromosome 11, clone RP11-583f24, complete sequence.		
ACCESSION	AC023078		
VERSION	AC023078.9	GI:20198710	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Choepeil,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeReteliano,K., Dewar,K., Domino,M., Doyle,M., Pensthor,J., Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J., Gardyla,S., Grant,G., Hagos,B., Heatford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Landers,J., Lehoczeky,J., Levine,R., Lien,C., Liu,G., Locke,K., Macdonald,P., Margis,N., McEwan,P., McChurt,A., McKernan,K., McPheters,R., Meldrim,D., Menus,L., Morrow,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Ollivar,T.M., Peterson,K., Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rothman,D., Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,D., Teefeay,S., Theodore,J., Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (08-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	3 (bases 1 to 163718)		
AUTHORS	Bitren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campiolo,A., Chang,J., Chazaro,B., Choepeil,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeReteliano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyla,S., Ginde,S., Gird,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,K., Karatas,A., Kellis,C., Larocque,K., Lamazates,R., Landers,T., Lehoczeky,J., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Margulis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Menus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,C., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roy,M.J., Rossetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strausen,N., Subramanian,A., Talamas,J., Teefeay,S., Theodore,J., Topham,K., Travers,M., Trivis,N., Trigliio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zaitoun,J., Zembek,L., Zimmer,A. and Zody,M.		

TITLE Direct Submission
 JOURNAL Submitted (19-APR-2002) Whitehead Institute/MIT Center for Genome
 RESEARCH Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 163718)
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukagaler, B., Brown, A., Camarata, J., Campolano, A., Chang, J.,
 Chazaro, B., Choquet, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., Deatello, K., Dewar, K., Diaz, J. S., Dodge, S.,
 Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Katat, A., Karatas, A., Kelle, C., Larocque, K., Lamazares, R.,
 Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
 Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Riedack, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Straus, N., Subramanian, A., Talamas, J., Teefaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigliolo, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
 JOURNAL Submitted (23-APR-2002) Whitehead Institute/MIT Center for Genome
 RESEARCH Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Apr 19, 2002 this sequence version replaced gi:15706197.
 COMMENT All repeats were identified using RepeatMasker:
 smhc, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WITR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submission@genome.wi.mit.edu
 ----- Project Information
 Center project name: L6254
 Center clone name: 583_F_24

FEATURES
 source

Location/Qualifiers
 1. 163718
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11"
 /clone="RP11-583F24"
 /clone_id="RPC1-11 Human Male BAC"
 complement(1359..1483)
 repeat_region
 1769..1864
 /rpt_family="MIR"
 repeat_region
 /rpt_family="L2"
 complement(2486..2593)
 repeat_region
 /rpt_family="MIR"
 3132..3212
 /rpt_family="L2"
 complement(3321..3328)
 repeat_region
 /rpt_family="MIR3"
 3523..3709
 /rpt_family="MLT1J"
 3815..4164
 /rpt_family="MER2"
 4460..4932
 /rpt_family="LTR33"
 5867..7374
 /rpt_family="L1PA4"
 complement(8866..8938)
 repeat_region
 /rpt_family="L1PB1"
 complement(9021..9603)
 repeat_region
 /rpt_family="L1PB1"

repeat_region 10764..10796
 /rpt_family="AT rich"
 repeat_region complement(11072..11270)
 /rpt_family="L2"
 repeat_region 1819..12173
 /rpt_family="MLT1B"
 complement(12553..12696)
 repeat_region
 13423..13787
 /rpt_family="L1MC4a"
 repeat_region 13788..13816
 /rpt_family="L1MC4a"
 repeat_region 13817..13965
 /rpt_family="L1MC4a"
 repeat_region 13966..14276
 /rpt_family="L1MC4a"
 repeat_region 14277..14341
 /rpt_family="L1MC4a"
 repeat_region 14358..14420
 /rpt_family="L1MC4a"
 repeat_region 14586..14659
 /rpt_family="L1MC4a"
 repeat_region 14586..14659
 /rpt_family="AT rich"
 repeat_region complement(14769..15433)
 /rpt_family="L2"
 15192..15199
 /note="<30 qual SNGL region"
 repeat_region 15437..16450
 /rpt_family="L1MD2"
 repeat_region complement(16453..16536)
 /rpt_family="L2"
 18608..18633
 /rpt_family="L1MD2"
 repeat_region complement(18931..19622)
 /rpt_family="L1PA8"
 19616..20496
 /rpt_family="L1PA8"
 repeat_region 20970..22109
 /rpt_family="L1P"
 repeat_region 20970..22109
 /rpt_family="L2"
 complement(22343..22686)
 repeat_region 24730..24895
 /rpt_family="L1MB4"
 complement(24730..24895)
 /rpt_family="MIR"
 25216..25285
 /rpt_family="L2"
 repeat_region 26057..26142
 /rpt_family="L2"
 repeat_region complement(26150..26307)
 /rpt_family="MIR3"
 26385..26640
 /rpt_family="MLT1J"
 repeat_region 26755..27105
 /rpt_family="MER2"
 27397..27858
 /rpt_family="L2"
 complement(28675..34751)
 repeat_region 34756..36869
 /rpt_family="L1PA9"
 36870..37829
 /rpt_family="L1PA8"
 repeat_region 37830..39164
 /rpt_family="L1PA8"
 complement(39165..39467)
 repeat_region 39468..40356
 /rpt_family="L1PA8"
 repeat_region 40350..40798
 /rpt_family="L1MD2"
 40872..41052
 /rpt_family="L1MD3"

Query Match 69.2%; Score 5966.2; DB 9; Length 163718;
 Best Local Similarity 84.4%; Pred. No. 0;

Matches 7537; Conservative 0; Mismatches 968; Indels 420; Gaps 57;

QY 1 TGTATGAAGCCATGTCACCTTATATCCAAAACGAGAAAGATATATCAAA--AAAGAA 58
 Db 117203 TCTATGAAGCCAGTGTCAACCTATATCAAAAACGAGAAAGATATATCAAA--AAAGAA 117144
 QY 59 ACTATGAGCCAGTACGATGATATATCATGACAGAAATCCCAACAAAATCTAGCTA 118
 Db 117143 ACTATGAGCCAGTACGATGATATATCATGACAGAAATCCCAACAAAATCTAGCTA 117084
 QY 119 ACCCAATCCACAGCATATCAAGAGATATCCACCATTTGTCAAGTGGTTTCATACCA 178
 Db 117083 ACCCAATCCACAGCATATCAAGAGATATCCACCATTTGTCAAGTGGTTTCATACCA 117024
 QY 179 GGGTGA--GGATAGGTTAATATATCAAGTCAATTAATGTGATACATCAATTAACGA 237
 Db 117023 GGGTGAAGGATGGGTTAATATATCAAGTCAATTAATGTGATACATCAATTAACGA 116964
 QY 238 TTAATAAACAAAATCATGATCATCTCAATAGATGTAAGAAAAGATTTGACAAATCT 237
 Db 116963 TTAATAAACAAAATTTTCATGATCATCTCAATAGATGTAAGAAAAGATTTGACAAATCT 116904
 QY 298 AACATTTCTTTATGATTAATAACCTTCAGCAAAATCGACATAGAAAGACATACCTTAAT 357
 Db 116903 AACATTCGGTATGATTAATAACCTTCAGCAAAATCGGACATAGAAAGACATATCTTAACA 116844
 QY 358 TTAATAAACCCATATATGACGACCCACAGCAAACTTATATCTGATGGGAAAGTTGA 417
 Db 116843 TAGTAATAAGCCACCGATGACAACTATATGCCCACAACTTATATATGGGAAAGTTGA 116784
 QY 418 AAAATTTGCTCCCTGAG-----AACTGAAACAAGACAAGATG-CTACTTTGACCACTT 469
 Db 116783 AAAATTTCTCCAGAGTTGAAAAATTTCTCCAGAAACAAAGATGCCCCACTTTGACCACTT 116724
 QY 470 CTATTCACATAGTATGAGAGTTTATGACGACATCAAGCAAGAAAGAAATCAAG 529
 Db 116723 CTGTTCAACATAGTATGAGAGTTTATGACGACATCAAGCAAGAAAGAAATCAAG 116664
 QY 530 GGCACCCAAATCAATTAAGAGAACTCAAACTGCTGTTCACTGATGATATGATTTGA 589
 Db 116663 GGCATCCAAATCAATTAAGAGAACTCAAACTGCTGTTCACTGATGATATGATTTGA 116604
 QY 590 TACCTAGAAAACCCCTAAGACATCATCAAGAAAGCTCTAGAACTGATACATAAATTCAGT 649
 Db 116603 TACCTAGAAAACCCCTAAGACATCATCAAGAAAGCTCTAGAACTGATACATAAATTCAGT 116544
 QY 650 AAAATTTGAGATACAACTAAATGTACAAATCAGTACACTGCTATACACCAACAGT 709
 Db 116543 AAAATTTGAGGATACAACTAAATGTACAAATCAGTACACTGCTATACACCAATAGT 116486
 QY 710 GACCAAGCTGAGATATCAATCAAGAACTCAAACTTTTCAATAGCTGTAAAAAAA--- 766
 Db 116485 GATCAAGCTGAGATATCAATCAAGAACTCAAACTTTTCAATAGCTGTAAAAAAA--- 116426
 QY 767 -----TACTTAAGATATTTCTTACCAAGAGAGTAAAGACCTCTACAAAGA 813
 Db 116425 AAAAAAACAACCTTACAGATATATCTTACCAAGAGAGTAAAGACCTCTACAAAGA 116366
 QY 814 AAACCTACAAAACAGCTGACATCATAGATGACAAACAAAGTGAACACATCCATGTC 873
 Db 116365 AA-----CTATGAAACCTGTAATCATATAGATGACAAACAAAGTGAACACATCCATGTC 116309
 QY 874 TCAATGATGGTATGATATCAATATTTGTGAAAATGACCAATTTGGCAAAAGCAATCTACAG 933
 Db 116308 TCAATGATGGTATGATATCAATATTTGTGAAAATGACCAATTTGGCAAAAGCAATCTACAG 116249
 QY 934 TTAATGCAATTTCCACAAAATATATCATCTTTCTTACAGAACTAGAAAACAAAT 993
 Db 116248 TTAATGCAATTTCCACAAAATATATCATCTTTCTTACAGAACTAGAAAACAAAT 116190
 QY 994 TCTAAATTCATATGGAACAACAACAAAACAAACAAACAAACAAACAAACAAACAAACAA 1053
 Db 116189 TCTAAATTCATAT-----ACAAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 116145

QY 1054 AGACTTAGCAAAAAGAACAAATCTGAGGACATCACTTACCATCTTCAAACTATCTAC 1113
 Db 116144 ACCTTAGCAAAAAGAACAAATCTGAGGACATCACTTACCATCTTCAAACTATCTAC 116085
 QY 1114 AAGCTATATATCACCAAAACATCATGACATGACATTAACCTAGGACATAGACCAATGG 1173
 Db 116084 AAGCTATATATCACCAAAACATCATGACATGACATTAACCTAGGACATAGACCAATGG 116025
 QY 1174 AAAAGAGAGAAATCCAGAAATTAAGCCAAATATATATAGCAACTGATTTTGACAA 1233
 Db 116024 AAAAGAAATGAAATCCAGAAATGAAACCAATATCTTACAGTGTACTGATTTTGACAA 115965
 QY 1234 GCAACAAAACAAATTAAGTGGGAAAAGACATTTAGTTAAACAATTTGGTGTGATAT 1293
 Db 115964 GCAACAAAACAAATTAAGTGGGAAAAGACATTTAGTTAAACAATTTGGTGTGATAT 115905
 QY 1294 TGGCAAGCCACATGAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 1353
 Db 115904 TGGCAAGCCACATGAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAA 115846
 QY 1354 TGTATCAAGATGATCAAAAGACTTAAATCTGAGCTTAAACCATTAATTTTGAAAG 1413
 Db 115845 CAATCAAGATGATCAAAAGACTTAAATTAAGACCTTAAGCAATTAATTTTGAAAG 115786
 QY 1414 TAAATCAGAAAATGCTCTTCAAGATTCATCACTTAAAGCAAGCTTCAATGCAAGAACCC 1473
 Db 115785 TAAATCAGAAAATGCTCTTCAAGATTCATCACTTAAAGCAAGCTTCAATGCAAGAACCC 115726
 QY 1474 AAAAGTAAATGCAACAAAACAAATTAATATGATGAGCTTAATTAATTAATTAATTAAT 1533
 Db 115725 AAAAGTAAATGCAACAAAACAAATTAATATGATGAGCTTAATTAATTAATTAATTAAT 115666
 QY 1534 TTGCGAGCAAAAACAAATATATGACAGCAACAGCAACCCACGAGTGAAGAAAT 1593
 Db 115665 TTGCGAGCAAAAACAAATATATGACAGCAACAGCAACCCACGAGTGAAGAAAT 115606
 QY 1594 CTTCAAACTAAGCACTGACCTAAGAACTAATATCCGGAATCCCAAGAACTCAACA 1653
 Db 115605 CTTCAAACTAAGCACTGACCTAAGAACTAATATCCGGAATCCCAAGAACTCAACA 115546
 QY 1654 AATCAGCAAGAAAGAAACAAACATCCATGAAAGTGGGCTAAGACATGATATGACA 1713
 Db 115545 AATCAGCAAGAAAGAAACAAACATCCATGAAAGTGGGCTAAGACATGATATGACA 115490
 QY 1714 ATTCTCAAAAGAAATATCAATATGCGCAACAAAC--AGSAAAAATGCTTAATCACT 1771
 Db 115489 ATTCTCAAAAGAAATATCAATATGCGCAACAAAC--AGSAAAAATGCTTAATCACT 115430
 QY 1772 AATGATTAAGGAAATGTAATCAACCTGTAATGCAATCACTTAATCTCTGCAAGAT 1831
 Db 115429 AATGATTAAGGAAATGTAATCAACCTGTAATGCAATCACTTAATCTCTGCAAGAT 115370
 QY 1832 GGTATATTTTAAATCTTAAATTAATATGATGTTGGTGGTCTGTGTGATTAAGAAC 1891
 Db 115369 GGTATATTTTAAATCTTAAATTAATATGATGTTGGTGGTCTGTGTGATTAAGAAC 115310
 QY 1892 ACTTTTACCTGTGTGGGAAATGTAATCTTACGCAACATATGAGAAACAGTGTGAA 1951
 Db 115309 ACTTTTACCTGTGTGGGAAATGTAATCTTACGCAACATATGAGAAACAGTGTGAA 115250
 QY 1952 ATTTCTTAAAGAACTAAGATGACATTTTGTCCAGCAATCCCA----- 2000
 Db 115249 ATTTCTTAAAGAACTAAGATGACATTTTGTCCAGCAATCCCA----- 115190
 QY 2001 ----- 2000
 Db 115189 CTACCCAGAGAAAGAAAGATCATTTTACGAAAAAGCACTTGCATGTGATTTATAGC 115130
 QY 2001 ----- 2000
 Db 115129 AGCAATTTGCAATTAACAAAATATGGAACAGCTCAAAATGCCATCAATCAACAAATGG 115070

[illegible]

Db	113994	ATCCAGAAAGGCCCTCCACCTTTACACAGAGGAGATCTTACCTCTTGACAAAAGAA	113935
Qy	3051	AAGGAACGATAAGGGTAAGTACCAAGAACTCTTCTTCCACAGTCAGTTATGATTTTG	3110
Db	113934	AAGGAAAGCAAGGGTAAAGTACCAAGAACTGCTCTTTATATAGTCAGTTATGATTTTG	113875
Qy	3111	CTGTAAAGTCAATGCTTATGCTTCACTTGTGTACATGCAAGGGGCTCACAGCTTG	3170
Db	113874	CTACAAAGATCACTTCTCTGTGTTCACCTCCGTCTACATCAGGGGATCATGATCTTG	113815
Qy	3171	TTTCAGAAAGACAGAGACATGAAGCTCTCTTCAAGAACTGAGTGCTGCAACCCAA	3230
Db	113814	TTTCATGAAAGACAGAGACATGAGGCTCTCTCTGAAACCGATGTGTACAGCCAG	113755
Qy	3231	ACTGTGTAGGCTCTAAATGTGTGTCCCTCTTAATTTATCTCCCATATCACTCTTC	3290
Db	113754	TCTGTGTAGTTCCAAAATAGGGTCCCCGCTTGTATTTATCTCCCATATCACTCTTT	113695
Qy	3291	ATTCGAATCATTCAAATCTGTCCCTCATGAGAGACTGTGCTCTTACATTCATTTAACA	3350
Db	113694	ATTCGAATCATTCATCTGTCCCTCATGAGAGAAATGCTACTTCAATTCATTTAATGA	113635
Qy	3351	GGAAGGGGACATGACAGGCAATTTCTTCCAGATTTAACTGATAG - AACCAATTTCTT	3409
Db	113634	GGAAGGGGACATGACAGGCTGTGCTTCTTACAGCTAATCTGTGAAACCAAGTTTGCT	113575
Qy	3410	GTTTCACTTACTTTTCAAAATTTATCTTCTTGTGCTATCTGAAAGGCTTAAAGAAATA	3469
Db	113574	GTTTATCTTACTTTTCAAAATTTTCTTCTTGTGCTATCTGAAAGTTTCAAGAAAGACA	113515
Qy	3470	TAGATGGCCC-----ATAATTAAGAGTGTTCATGAGAAAGTATTTACAAAGATG	3522
Db	113514	TGGATGTGTCCAAACTGAATAGATTTAGAGAGTGTTCATGAGGAAAGTATTTACAAAGACA	113455
Qy	3523	CACAGAGTTAA-----GGGTACAGATCTTAAAGCAGCAATACATAGGGG	3565
Db	113454	TGCAGAGGTTAAAGAAAGAAAGAAATGATCATGGGTACACACCTTAAGCAGCAATACATAGGGG	113395
Qy	3566	AGCATTACTTCTCCCTTAGGCTGAAACGGAACAGGAAAGAGACAGTTACATTTGTC-----	3621
Db	113394	AGCATTACTTCTCTCCCTTAGCTGAAAGGGAACAGGAAAGAGACAGTTATTTATTTGTCACCA	113335
Qy	3622	--GCCATAGCCATAGCTGTAGCCATAGAGGTGGAGAGACATAGCAGGCAAGTGGAGAAAG	3679
Db	113334	TAGCATATGCCATATAGCTGTATCATATAGGGGTGGAGAAACAAAGCGGCGAGGTGGAGAAAG	113275
Qy	3680	CCCTGCGTGGCCACGACACGACCAACAGGCTGATATAGTTGGATCTGTGTCCACCA	3739
Db	113274	CCCTGTATGGCCACACACAGGCAACAGGCTGATATGCTTGTGATCTGTGTCTTCACCC	113215
Qy	3740	AAATCTCATGTGTATTTGTAATTTCCAAATGTGGAGAAAGGGCTGTGGAGATGATTAAT	3799
Db	113214	AAATCTCATGTGTATTTGTAATTTCCGATGTGGAGAAAGGGCTGTGGAGATGATTAAT	113155
Qy	3800	TAGATCAGCGGGATGTTTTCGATGAATGTTTAAACCAATCCCTTGTGTATTTGTTGT	3859
Db	113154	TAGATCAGCGGGATGTTTTCGATGAATGTTTAAACCAATCCCTTGTGTATTTGTTGT	113095
Qy	3860	TGTGATCTGACGAGTTCATGAATACTGTTTAAAGCGTGTAGCACTCTCTCC- 3918	
Db	113094	TGTGACACTGATGAATTTCTCATGATATCTAGTTTAAAGATGTGTATCACTCTCTCC	113035
Qy	3919	-CTCTCTCTTACTCGGCTCTACATATGAAAGGCTCGCTCCCTTGTGCTTTACAC 3977	
Db	113034	TCTCTCTCTGCTCTGCTTTCACATATGAAATGCTCTCTCTCCCTTGTGCTTTACAC 112975	
Qy	3978	AGGATTTGAAGCTTCTCTAGAGGCTCTCCCAAGAAAGGCTGTATGCTTCTTGTATAGT 4037	
Db	112974	ATGATATGAAGCTCTCTAGAGGCTCTCTTAAAGCAAGAGCTGTATGCTTCTGTATAG 112915	
Qy	4038	CTGTAGAGCTATTTAGCCAGTTAAACCAATTTCTTCAATTAATTTCCAGTCTCAGTAA 4097	

Db	112914	CTGCAGAGCACTGAAACCAATTAAACCTGTTTCTTCTATAAATTTCCCACTTCAGGATAT	112985
Qy	4098	TCTTTTATGCAATTTGAGATGAACTAATATACACAGACAGAGCCAGAGATGGAATCC	4157
Db	112854	TCTTTGTAGCAATTCAGAAATGAACATAATATACCTGGACAGAGGCCAGAGGTGGAATCC	112795
Qy	4158	CAAGGTGCTTCTCGCTGCTCCAGCTCTCCGTGCTGCTGCTCCAGTGTCAATTCAC	4217
Db	112794	CAAGGTGCTCTCTGCTGCTCCAGCTCTCCGTGCTGCTCCAGTGTCTCAATTCAC	112735
Qy	4218	CCAGAAACCAAGAAATPAAAGAATCCCACTGATGTGTACATAGAACCACTCTCTGGG	4277
Db	112734	CCAGAAACCAAGAAATPAAAGAATCCCACTGATGTGTACATAGAACCACTCTCTGGG	112675
Qy	4278	ATGTCAAAACGAGATPAAAGAATGGAAGCAAAATCTCATGTGTAAATGACATACCT	4337
Db	112674	ATGTCTAACAGAGATAGAGAAGAAATGGAAGCAAAATCTCTGGCACTAAAGATACCT	112616
Qy	4338	CTCACCCTTCTGTATCTCTCTAATTCCTGGGGCTTCTCTATC-----TGATTTGATCCC	4391
Db	112615	CACACCTCTGTATCTCTCTAATTCCTCTGCTGTGTTTTCTGTAAAGGTATGTATCT	112566
Qy	4392	TGTCATATTCAGCTCTATCAGACTCTTAAATGTGTGCTGTCTTCTCTACTCTCAC	4451
Db	112555	GGTCTCATTTCAACTCTATCAGAC-----TAAATTTGGCAATCTCTCTACTCTCAC	112501
Qy	4452	TTTTATGCAAAATGTTTGCATTTGTTAAAAATGCATAGAAAATAAATGTAATTTTAA	4511
Db	112500	TTCTATGCAAGATGTTTGCATTTGTTAAAAATGCATAGAAAATAAATGTAATTTTAA	112441
Qy	4512	AAGA-----ACATATGATTTTGTATGAAATATAAGTTTGGCTGTCTCATPAAAGACAG	4567
Db	112440	AAGAGAGGACATATTTCTTTTGTATGAAATATAAGTTTGGCTGTCTCATPAAAGACAG	112381
Qy	4568	AGAGAAATATCTTAAACCAAGAAATATATGTTGTGCTCTGGGCTCACTAGGTTCTGATC	4627
Db	112380	AGTAAATATATCTTAAACCAAGAAATATATGTTGTGCTCTGGGCTCACTAGGTTCTGATC	112322
Qy	4628	TACAGATTCAAACAACTACAGAGAGAACTTTTCCAAAATAAAGTGTGGCCGAGTTGT	4687
Db	112321	TGCAGATTCAAACAACTACAGAGAGAACTTTTCCAAAATAAAGTGTGGCCGAGTTGT	112262
Qy	4688	GTAATGTACTGAACAGGTACAACTTGTATTTCTTTGTCAATATTTCTGAAAACTACAT	4747
Db	112261	GTAATGTACTGAACAGGTATAACTTGTATTTCTTTGTCAATATTTCTGAAATGATGAGT	112202
Qy	4748	-ATAACAACAACCTATATATGCAATTTGCAATTTGTCAATTTCTGAAAACTT-ATAATG	4805
Db	112201	AAAAAAAACAATTTATATATGCAATTTGCAATTTGTCAATTTCTGAAATGTAATGAAAT	112142
Qy	4806	ATTTATGATCTGGAGAGAAAGTGCATAGATATATACAAATACAAAT-----ATTA	4858
Db	112141	ATTTATGATCTAGAGAGAAAGTGCATAGATATATACAAATATACATCGTTTGTATA	112082
Qy	4859	GGAATTTGAGCATCTGCAGATTTTGGTCTGTGCTGGGGTTCTGAAAAAATCCCTGTATA	4918
Db	112081	AAGAATTTGAGCATCTGCATATTTTGGTGTGTGCTGTGCGAGTTGCGAGAACCAAGCCCAATGA	112022
Qy	4919	ATACACAAAATATGACACTCTGCAAGATCTGAATAGAAAGTCCAAAGCAATCATATTCG	4978
Db	112021	ATACACAGATATGACACTCTCTAGATCTCCAGCTGAAAGTCCAAAGCAATCATATTCAG	111962
Qy	4979	AATTCAAAAAATGTGCTGCCAGATTCCTAGAGAGTGTGCTCATCTGTGATCTTAC	5038
Db	111961	AATTC-AAAAATGAAACTGCCAGATGCTCTAGAGATGTGCTCATTTTGTGATCTGCTG	111903
Qy	5039	ATGTTTCCAGCAGCATTAAGCATTCAGTCTTATGAAAAAAGACAGAGGGAAGAGAG	5098
Db	111902	ATGATTTCCAGCAGCATTCAGCATTCAGCATTCAG-AAAAAGATGAAAGAGAGAGAG	111844
Qy	5099	CTTTGCTCTCTTATTAATTCATAGAGCCAGAGACTGCTTCTGTCACTTTTGTGATTTCT	5158
Db	111843	CTTTGCTCTCTTATTAATTCATAGAGCCAGG-----GATCACTTTTGTGCACTTT	111793

QY	5159	CCACTTAACAGACCCCTGCTCATGGGATGTGATCCAGATCAAGGAAACCTGGATGTGGG	5218
Db	111792	ACATGTAAGGCTCTGCTCATATATATGCA-CAATCATCAAGAACCTGGATGTGGG	111734
QY	5219	TCCTTGCTGCTTGTACATTTCTCAGAAAGGTTATGTGACCAAAAAGAAATCTT----	5274
Db	111733	TCCTTGCTGCTTTTACCTCTCAGAACTGTTATGTGACCAAAAGAAATGATTAAT	111674
QY	5275	-GGGGCAACAGAGAGCTCTTGAGCCCGAGCTGCTCTGATTCGTGCTCAACATCAAG	5333
Db	111673	GGGGCAACCTGAGAGCTCTTGAGCCCGAGCTGCTCTGATTCGTGCTCAACATCAAG	111614
QY	5334	TTTTTCAGNACCTCCAGAAAATTAATTAATGTTGGGCGAGAAACAGAACTGAGTCTCG	5393
Db	111613	TTTTTCAGNACCTCCAGAAAATTAATTAATGTTGGGCGAGAAACAGAACTGAGTCTCG	111554
QY	5394	TGACAGACTCCAGGAGCACGGGGCTGGTATTTGAGACTGCTCTTCAATGTTGTGAAC-CAGG	5452
Db	111553	AACAGGGCTCCAGGGGACAGGGACGTGGATTTGAGACTGCTCTTCAATGTTGTGAAC	111494
QY	5453	AAAAACCTTAATTTCTAGAGGCTTAGGCTTCATCTATATGTAATAGATAATATACATA	5512
Db	111493	AAAAACCTTAATTTCTAGAGGCTTAGGCTTCATCTATATGTAATAGATAATATACATA	111434
QY	5513	GACAGCTTTTAAAGAACATCATGAGATGTTAAACAATGCTAAATGTTGGTATACAC	5572
Db	111433	GATGCTTTTAA-----GAAAGACATGCTAAATGTTGATGATATAC	111392
QY	5573	AGTGAATAAACAAGGATGATCTTACCTTACGATCTTGCGGGTTTATGAGGAAGCAA	5632
Db	111391	AGTGAATAAACAAG-----CTTATGGAGCTTCAAGGTTTCAATGAGGAAGTTAA	111345
QY	5633	ACATATCATACCATACCTATAGATGAGACAACAAGTTTATGTCCTGAGTGTGATAACAG	5692
Db	111344	ATATATCATACCTTAATATATATAGTGGACAAGATGTTTATGTCCTGATGTGATAACAG	111285
QY	5693	AGGTTCTCTTTTCTCCCATTTTCTTTTGGGCGAATCAAGCTGTGGCAGCTGTCTC	5752
Db	111284	ATGTTCTCTTTTCTCCCATTTTCTTTTGGGCGAATCAAGATGTTTATGTCCTGATGTGATAACAG	111225
QY	5753	CCTAAGAGAGCTCATGATGATGACACTACTCTGATGCTCTCTATATCTCCAGAGAG	5812
Db	111224	TCCTACTAAGAGTCAAGATGATGATACATTAATCTTGATGTTTCTCTGATCTCCACACAG	111155
QY	5813	GATGATCTTCTTTCCACTCGAGAGACTCTGCCCATGATGATTTCTTGGATTCAGAGC	5872
Db	111164	GGTGAATCTTCTTTCCACTCGAGAGACTCTGCCCATGATGATTTCTTGGGATTCAGAGC	111105
QY	5873	AAAGGTGGGCTCTGATATGGCAAAAAGAACTCTGAATTTGTCCTAAATGGACGCACT	5932
Db	111104	AAAGGTGGGCTCTGAGTATGGCAAAAAGAACTCTGAATTTGTCCTAAATGGATTTCACT	111045
QY	5933	CACCTCTAT-TTTTCCCTATTTTCAATTTGCTTCTCATTTCTCATCTGAGATTTGTTAGG	5991
Db	111044	CTCCTCTATATTTGTCCTATTTGATTTAGCTTCTCATTTCTCATCTGAGATTTGTTAGG	110985
QY	5992	TTAATTTTTTTTTTCAGCCCAATTTTACGTCAACTTGGATTTAACTTGAATCAC	6051
Db	110984	TTAATTTTTTTTTTCAGCCCAATTTTACGTCAACTTGGATTTAACTTGAATCAC	110925
QY	6052	TCCCTACTTACCCCCCTTAAACATGTATATATGACACATATGATGTGGTGGCTCAAG	6111
Db	110924	TCCTGTGCTTTA-CTCCCTCTAAACATGTATATATGACACATATGATGTGGTGGTCAAG	110866
QY	6112	GGCTGTGAAAAAATGATCATGAGTCAAGCCCTGCTGGGCTCAACTTCATCTATATAT	6171
Db	110865	GGCTATGCTCAAAAGCTGATCATGAGTCAAGCCCTGCTGGGCTCAACTTCATAC-----T	110813
QY	6172	ATATATACCCCCCGAGAAATATATATCTCTTTATATCTTAATTTCAATTTCTGCATA	6231
Db	110812	ATATATGCTCCAGGACAAATATATTAATTTCTTTTGTGTCTC-CAGTGCATTTATGGCATA	110755

QY 6232 CAGGAATATACTAATTTTAACTCTAGGCTCTTCAGATGATTTAAAGAGCAATACCT 6291
 DB 110753 TGGGAATAAATACATATCTTAACCTCCTAGGCTCTTCAGATTAATTAAGAGCAATACCT 110694
 QY 6292 AATAAATCTCAATCAGCTGCTGTATTTCTCCAAATTAGACCTAATCTCATCTCCAG 6351
 DB 110693 AATPAATCTCAATCAGCTGCTGTATTTCTCCAAATTAGACCTAATCTCATCTCCAG 110634
 QY 6352 TTAGAATTTGAGAAATATCTCTCTTAAACAACCAAGCCCTACCTTCTCTAATTTCCAC 6411
 DB 110633 CTGAATCTGCAATCTCTCTCTTATATCTCAAGCCCTATCTAGCCCTAATTTCC-C 110575
 QY 6412 TCATGACCTCTCTCAATCAAAATGTTGACCAAAAGAGCTAATCAAAATCTCCC 6471
 DB 110574 TCATGACCTCTCTCAATCAAAATGTTGACCAAAAGAGCTAATCAAAATCTCCC 110516
 QY 6472 GAAAGAGAGAAATAGATAGTTTACATTTGTATTAATCTCAGAGAAACCTTAGTATGCCCC 6531
 DB 110515 GAAAGAGAGAGAAATAGATAGTTTACATTTGTATTAATCTCAGAGAAACCTTAGTATGCCCC 110456
 QY 6532 CATACATATTTCCACACTTCAATTAACCTGACAGTGCATCTAGGCTCAACCTCACTTA 6591
 DB 110455 CATACATATTTCCACACTTCAATTAACCTGACAGTGCATCTAGGCTCAACCTCACTTA 110396
 QY 6592 CTCTTCTCTGCTCTATGCTGAGCAATTAAGCTCAGACCCCAACCTTACCCTCAACT 6651
 DB 110395 CTCTTCTCTGCTCTATGCTGAGCAATTAAGCTCAGACCCCAACCTTACCCTCAACT 110336
 QY 6652 GTGAC-AAAATGCTTCTAGGGGTTGGGCAAGCACTGAGCTTATTTTAAAGCA 6710
 DB 110335 GTGACAAAATGCTTCTAGGGGTTGGGCAAGCACTGAGCTTATTTTAAAGCA 110276
 QY 6711 CATCAGTGTCAATTTAGGTTTGGGCACTCAATCAATCTTCTTCAACAGAGATAGA 6770
 DB 110275 CATCAGTGTCAATTTAGGTTTGGGCACTCAATCAATCTTCTTCAACAGAGATAGA 110216
 QY 6771 GCTGTCCAAATGATTTCTGATGAATGAATTTTCTTCACTTAATATGCTGTGT 6830
 DB 110215 GCTGTCCAAATGATTTCTGATGAATGAATTTTCTTCACTTAATATGCTGTGT 110156
 QY 6831 TCTAATGCTTCAATGCTTCAATTTTATTTTCAATTTCAATCAATCAATCAATGCT 6890
 DB 110155 TCTAATGCTTCAATGCTTCAATTTTATTTTCAATTTTCAATCAATCAATCAATGCT 110096
 QY 6891 CATTAAGCTTCTCAATGATCTTCAATGATGAATGAATGCTTATGAAAGCAATGCT 6950
 DB 110095 CATTAAGCTTCTCAATGATCTTCAATGATGAATGAATGCTTATGAAAGCAATGCT 110036
 QY 6951 GCTGTCTATG-ATAGGCACTAGAGATTAATTAATGTAATGTAATGCTGTCTGCAATG 7009
 DB 110035 GCTGTCTATGAAATAGGCACTAGAGATTAATTAATGTAATGTAATGCTGTCTGCAATG 109976
 QY 7010 ACTGACACATGATTTTCTGACCCAGGCTCCGCCATTTTCAAC--ATCCTAGC 7067
 DB 109975 ACTGACACATGATTTTCTGACCCAGGCTCCGCCATTTTCAACCTTATCTAGC 109916
 QY 7068 GAAGATCCATTTTCTCTGCTTCAATGATGATCTTTTCTGCTGACAGATGACC 7127
 DB 109915 AAAGGTCTCATTTTCTCTGCTTCAATGATGATGCTTTTCTGCTGATGATGACC 109856
 QY 7128 AGTCTGTGATGAGGCTGACCAACCACTTTTGTATGTAATTTCTTCCACTGAG 7187
 DB 109855 AGTCTGTGATGAGGCTGACCAACCACTTTTGTATGTAATTTCTTCCACTGAG 109796
 QY 7188 AGAATAATTTGAGGCGAGATAGATATCATGGGCTCCAGAGACTGGCTAGATGAGT 7247
 DB 109795 AG-AAAATTTGAGGCGAGATAGATATCATGGGCTCCAGAGACTGGCTAGATG 109737
 QY 7248 GAGGCTGTTTGTATCTTAATGTTATCCCATGTCAGCAGAACTTGTGTGCACTAGAGA 7307
 DB 109736 GAGGCTGTTTGTATCTTAATGTTATCCCATGTCAGTACAGAACTTGTGTGCACTAAGAGA 109677
 QY 7308 GAGGTACGCTTCAAGTCAACAAGAACTGATTTCAAACTGATTTGAGAGACCCCAACC 7367

DB 109676 GAGGTACGCTTCAAGTCAACAAGAACTGATTTTCAAA-TGAAATGAGAGACCCCAACC 109618
 QY 7368 TTTTGAATGAGTAACTTATCTGTCGAGTCTCTG-ATCTCTCTCTTT-AAATAGAGACA 7425
 DB 109617 TTTTGAATGAGTAACTTATCTGTCGAGTCTCTGTCCTCTCTTTAAATGAGAGACA 109558
 QY 7426 GTAATCTCAATGAGGCTGTGGGAGAAATCAGAGATCAAAACGCTGTATACAT 7485
 DB 109557 GTAATCTCAATGAGGCTGTGGGAGAAATCAGAGATCAAAACGCTGTATACAT 109498
 QY 7486 CTGCTTCTGTTTCCAGGCTATCAGACTGGGCTTCTGAGCAATGATTCAAACATCCCA 7545
 DB 109497 CTGCTTCTGTTTCCAGGCTATCAGACTGGGCTTCTGAGCAATGATTCAAACATCCCA 109438
 QY 7546 GTCTTGGGTACAGAACTGACCAACCAATCAAGAGCTGAGAGACTCTCTTCAAGAG 7605
 DB 109437 GCCTTGATGAGAACTGACCAACCAATTAAGAACTGAGAGACTCTCTTCAAGAG 109378
 QY 7606 ACCCTGAGCTTCAAGGAGCTGACGTCGATGCTTCCCTGTGCGGCTGACAGAAACGG 7665
 DB 109377 ACCCTGAGCTTCAAGGAGCTGACGTCGATGCTTCCCTGTGCGGCTGACAGAAACGG 109318
 QY 7666 GTTGTGCTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7725
 DB 109317 GTTGTGCTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 109258
 QY 7726 AACCTGTCGAGGCGGACTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 7785
 DB 109257 AACCTGTCGAGGCGGACTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 109198
 QY 7786 CTCAATCAATTCGCAATTCCTCAATTCCTCAATTCCTCAATTCCTCAATTCCTCAAT 7845
 DB 109197 CTCAATCAATTCGCAATTCCTCAATTCCTCAATTCCTCAATTCCTCAATTCCTCAAT 109138
 QY 7846 TTTTAAAGGCTTAAGAGCTGAGGCGCATGAGCAACGAGGCTGCTGCTGCTGCTGCTG 7905
 DB 109137 TTTTAAAGGCTTAAGAGCTGAGGCGCATGAGCAACGAGGCTGCTGCTGCTGCTGCTG 109078
 QY 7906 CCATCTGTGATACCACTGCGC--CGCCCAAGATCTGTCAATGCTGATGCTGCTGCTG 7963
 DB 109077 CCATCTGTGATACCACTGCGC--CGCCCAAGATCTGTCAATGCTGATGCTGCTGCTG 109018
 QY 7964 TCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8023
 DB 109017 TCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 108958
 QY 8024 GTGTGCTGATTTCTGTTGGTGAACGTCAGATTTCAATTAAGTGGCGCTGCTGCTG 8083
 DB 108957 GTGTGCTGATTTCTGTTGGTGAACGTCAGATTTCAATTAAGTGGCGCTGCTGCTG 108898
 QY 8084 TTTTATGTGATTTCTGTTGGTGAACGTCAGATTTCTGTTGGTGAACGTCAGATTTCTGTT 8143
 DB 108897 TTTTATGTGATTTCTGTTGGTGAACGTCAGATTTCTGTTGGTGAACGTCAGATTTCTGTT 108838
 QY 8144 CCGGAAAGATGCGGCTGACCAAGCTGTGACATCTTCTTCAAGTCTGCTGCTGCTGCT 8203
 DB 108837 CCGGAAAGATGCGGCTGACCAAGCTGTGACATCTTCTTCAAGTCTGCTGCTGCTGCTGCT 108778
 QY 8204 TCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8263
 DB 108777 TCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 108718
 QY 8264 GAAAGCTTATTTTGTGATGATGATGATTTTCAATTTTCTGCTGCTGCTTAAAGACA 8323
 DB 108717 TGAAGTCAATTTTGTGATGATGATGATTTTCAATTTTCTGCTGCTGCTTAAAGACA 108658
 QY 8324 GTGCAACCCCATCATATTTCTTCTGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCT 8383
 DB 108657 GTGCAACCCCATCATATTTCTTCTGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCT 108598
 QY 8384 ACCTGAAGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 8443

Db 108597 ACTGAGCTGGTCTCCAGAGGCTCTGACAGCATGCTGAGG-----TGAAGTGGAG 108542
 QY 8444 GGTGGCTTCTCAGAGAACTTGGAGCTGTGGGAAAGCAGATTGAGCAGTGAAGAA 8503
 Db 108541 GGGCGCTTCTGAGGAAACCTGGAGAGCTCGGGAAGCAGATTCCGGCAGTGAAGAA 108482
 QY 8504 CCTCGGCTCTCTGAGAGAGATTGAGAGCAATGCTGCGCCAGCAGCTTGACAA 8559
 Db 108481 CCTTGGCTCTGAGTCAAGCAGGAGCTTGAAGAGCAATCTGCTGCGCAGCTTGACAA 108422
 QY 8560 TTATATGATTTTCTTACCTTCTGCTCAGAAATGCTCAAGG 8604
 Db 108421 TTACATGC-GTTTCTTACTGTTGTGCTCAGAAATGCTCAGCG 108378
 RESULT 5
 AC139482 165490 bp DNA linear HTG 04-FEB-2003
 LOCUS Homo sapiens chromosome 5 clone RP11-583P24, WORKING DRAFT
 DEFINITION SEQUENCE, 12 unordered pieces.
 AC139482 165490 bp DNA linear HTG 04-FEB-2003
 AC139482.1 GI:28201512
 HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 165490)
 DOE Joint Genome Institute.
 Sequencing of Human Chromosome 5
 Unpublished
 2 (bases 1 to 165490)
 DOE Joint Genome Institute.
 Direct Submission
 Submitted (04-FEB-2003) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: http://www.jgi.doe.gov

 Project Information
 Center Project Name: 1518520
 Center clone name: RPCT-11_583P24

 Summary Statistics
 Consensus quality: 147035 bases at least Q40
 Consensus quality: 155210 bases at least Q30
 Consensus quality: 158567 bases at least Q20
 Estimated insert size: 175000; agarose-fp estimation
 Estimated insert size: 164390; sum-of-coverage estimation
 Quality coverage: 3.94 in Q20 bases; agarose-fp estimation
 Quality coverage: 4.19 in Q20 bases; sum-of-coverage estimation
 NOTE: This is a 'working draft' sequence. It currently
 consists of 12 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 1 1249: contig of 1249 bp in length
 * 1250 1349: gap of unknown length
 * 1350 3680: contig of 2331 bp in length
 * 3681 3780: gap of unknown length
 * 3781 6668: contig of 2888 bp in length
 * 6669 6768: gap of unknown length
 * 6769 10866: contig of 4098 bp in length
 * 10867 10966: gap of unknown length
 * 10967 17019: contig of 6053 bp in length
 * 17020 17119: gap of unknown length
 * 17120 22433: contig of 5314 bp in length
 * 22434 22533: gap of unknown length

* 22534 34645: contig of 12112 bp in length
 * 34646 34745: gap of unknown length
 * 34746 49173: contig of 14428 bp in length
 * 49174 49273: gap of unknown length
 * 49274 67058: contig of 17785 bp in length
 * 67059 67159: gap of unknown length
 * 67159 84928: contig of 17770 bp in length
 * 84929 85029: gap of unknown length
 * 85029 114464: contig of 29436 bp in length
 * 114465 114564: gap of unknown length
 * 114565 165490: contig of 50926 bp in length.
 FEATURES
 source 1. 165490
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="RP11-583P24"
 /clone_1ib="RPCT human BAC library 11"
 ORIGIN
 Query Match 62.4%; Score 5380.4; DB 2; Length 165490;
 Best Local Similarity 84.0%; Pred. No. 0;
 Matches 6838; Conservative 0; Mismatches 906; Indels 394; Gaps 52;

QY 763 AAAATCTTAAGAAATTTCTTACCCAGAGGAGTGAAGGCTCTCAAGAGAACTACAA 822
 Db 114643 AAACCACTTAGAGGTTTAGCAAAAAGAGGTGAAGC-----TCTCAAGAAACTAT 114698
 QY 823 AACACAGCTGATCATATAGATGACACAAAGTGAAGAACATCCATGCTCATGATG 882
 Db 114699 GAACTGCTGAATCATATAGATGACACAAAGTGAAGAACATCCATGCTCATGATG 114758
 QY 883 GGTGAATCAATTTGTGAATTTGACCAATTTGCCAAAAGCAATTAAGTTCAATGCA 942
 Db 114759 GGTAAATCAATTTGTGAATTTGACCAATTTGCCAAAAGCAATTAAGTTCAATGCA 114818
 QY 943 ATTCCACCAAAATATCATCATCTTCTTCAAGAACTGAGAAAACATTTCTAAATT 1002
 Db 114819 ATTCCACCAAAATATCATCATCTTCTTCAAGAACTGAGAAAACATTTCTAAATT 114877
 QY 1003 CATATGAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 1062
 Db 114878 AAT-----ACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 114922
 QY 1063 AAAAAGCAATCTGAGGAGTCACTTACATCCATCTTCAACTATATCTAAGCTATA 1122
 Db 114923 AAAAAGCAATCTGAGGAGTCACTTACATCCATCTTCAACTATATCTAAGCTATA 114982
 QY 1123 ATACCAAAACATCATGAGTCACTTACATCCATCTTCAACTATATCTAAGCTATA 1182
 Db 114983 GTACCAAAACATCATGAGTCACTTACATCCATCTTCAACTATATCTAAGCTATA 115042
 QY 1183 GAGAAATCAAAATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1242
 Db 115043 GAGAAATCAAAATTAAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 115102
 QY 1243 AACTTAAGTGGGAAAAAGCAATTTAGTTAACAAATGTTGATGATTTATGGCAAGC 1302
 Db 115103 AACTTAAGTGGGAAAAAGCAATTTAGTTAACAAATGTTGATGATTTATGGCAAGC 115162
 QY 1303 ACATGGAAGAGATGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 1362
 Db 115163 ACATGGAAGAGATGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 115221
 QY 1363 ATGATCAAGAGCTTAATCTGAGACTTAAGCAATTAATTAATTAATTAATTAATTA 1422
 Db 115222 ATGATCAAGAGCTTAATCTGAGACTTAAGCAATTAATTAATTAATTAATTAATTA 115281
 QY 1423 AAAAATGCTTTAGATCACTTACAGCAAAAGCTTACATGAGCAAGAAACCCAAAG--TA 1480
 Db 115282 AAAAATGCTTTAGATCACTTACAGCAAAAGCTTACATGAGCAAGAAACCCAAAGGATA 115341

OY	1481	AATGCAACAAAAACAATAATTAATGATAGGACTTAATTAACATAAAGCTTTGCGA	1540
Db	115342	AATGCAACAAAAACAATAATTAATGATAGGACTTAATTAATTAAGCTTTGCGA	115401
OY	1541	GCAAAAAACATCATTTAGCAGACCAACACAACCCAGAGTGAAGAAATCTTCACA	1600
Db	115402	GCAAAAGAAATCATCATGACAGAACAAAGAAATTAACCACTGAAGTGGAGATATCTTCACA	115461
OY	1601	AACTAAGCATCTGACTAAGGACTAATATCTCGGAATCACAAGGAACCTCAACAAATCAGC	1660
Db	115462	AACTATGATCTGACAAAGGACAAATATCTCAGAACTCAAGAGAACTCAACAAATTAGC	115521
OY	1661	AAGAAAGAAAGCAAACTCCCACTGAAAGATGGCGTTAAGACATGAAATAGCAATTTCTCA	1720
Db	115522	AAGAAAGA---AAACAGTACATCAAAAAGTGGCGTTAAGGCAATGAATACAAATTTCTCA	115577
OY	1721	AAAGAAGATATCAAAATGCGCAACAAAC--AGGAAAAAATGCTTAAATCATCATATGATT	1778
Db	115578	AAAGAAGATATCAAAATGCGCAACAAACATTAAGAAAAATGCTTAAACCAAGATATCATC	115637
OY	1779	AGGAAATGCTAAATCAACCTGTAATGCATACCACTTAATCTCTGCAAGAAATGTCAT	1838
Db	115638	AGGAAACACAATCAATCAAAACCAATTAATGCATATGCACTTAACCTCTGAGAAATGGCCATA	115697
OY	1839	ATTTAAAAATCTAAAAATTAATAGATTTGTTGGTGTCTGTGGTATTAAGAAACATTTTA	1898
Db	115698	ATTTAAAAATCTAAAAATTAATAGATTTGTTGGTGTGTGGTGAAGAAAGTACATTTTA	115757
OY	1899	CAGCTCGTGGTGGAAATGTAACTGTGGCAACCATATGGAAGAAACAGTGGAAATTTCTT	1958
Db	115758	CAGCTCGTGGTGGAAATGTAACTGTGGCAACCATATGGAAGAAACAGTGGAAATTTCTT	115817
OY	1959	AAGGAATAAAAATGATGATGCAACATTTGATCCAGCAATCCCA-----	2000
Db	115818	AAAGAACTAGAAAGTATGATCTCAAAATTTGATCCAGTAATCCCACTCGGGTATCTAACCA	115877
OY	2001	-----	2000
Db	115878	GAGGAAAAAGAACTCAATTTACGAAAAAGACACTTGCACTTGATGTTATAGCAGACAA	115937
OY	2001	-----	2000
Db	115938	TTGCAATTAACAAAAATATGSAACAGCTCAAAATGCCATCAATCAACCAAGTGAATAAGA	115997
OY	2001	-----	2000
Db	115998	AAATTAATATATATATAGTGTGTATGTAATGTATGTATGTATATATATATATATATGTCT	116057
OY	2001	TTAAATATGTAATTAATATATATTT-----ATATACATGGAATACAA	2044
Db	116058	GTAATATATTAATATATATATATATATACACACACACACACATATACATGAAATATCTA	116117
OY	2045	CTCAGCCATAAAAAAGAAATTAATATGATGCACTTCAAGCAATTTAATAGGAATTGGAGAC	2104
Db	116118	CTCAGATTAATAAAAAGAAATGAATAATATATACATTTCAAGCAACCTAAGTGAATTTGAGAC	116177
OY	2105	CCTTATTTCTAAGTGGGGTAATCAGGAATGSAAAACCAAAATCATATATGTTCTCACTTAC	2164
Db	116178	CATTATTTCTAAGTGAAGTAATCTAGGAATGGAAGAACCAAAATCATATATGTTCTCACTTAT	116237
OY	2165	AAAGTGGGGCTAAGCTGTAGAGACAGAAAGCATAGAAATGATATTAATGAACTCTGGGGAC	2224
Db	116238	AAAGTGGAGCTAAGCTGTAGAGACAGAAAGCATAGAAATTAATCAATTAATGAACTTTGAGGAC	116297
OY	2225	TTGAGGGGAAAGATGGAAGAGAGGCGAGGATTAAGAATACCAATATGGGTACAGTGTAC	2284
Db	116298	TTGAGGGGAGTGAATGAAGAGAGATGAAGATTAAGAATACCAATATGGGTACAGTATAC	116357
OY	2285	ACTGCTCAGGTATAGGTGTCACAAAAATCTCAGAAATTAACCTAAGAACTTATATCATG	2344
Db	116358	ACTGCTTGGGTATAGGTGTCACAAAAATCTCAGAAATCACCACCTAAGAACTTATATCATG	116417
OY	2345	GAAGCAAAACCACTGTTCCCAAAATCCCAATGAAATTAATAATATATTAATGA	2404

Db	116418	TAAACCAACCAACCTGTTCCCAAAAACCCATTGAAATACA---AATAATATATATGA	116474
Qy	2405	TTTAATTTACAGAA-TTTTAAAAAGTTCACGTTCAGAGTTATATATATGAGTAA	2463
Db	116475	TTTGATTTACAGAACTTAAAAAAGTTAGCTTCAAGGTTGATG-AAAAGATGAAA	116522
Qy	2464	ATGAAAGGTGACGAGTGTGATGCTTCGACATGGGACTTAATTTTACCTTGCA	2523
Db	116533	ATGAAACGTGTAGCAAGTGTGATGCTTCAGCAATGGGATTTGACTTTCCTTGCA	116592
Qy	2524	AACCTTCCTGGCAATTGGAAAGAAAGTATACATGAAATATATACCACTATGATAAAG	2583
Db	116593	AACCTTCCTGACATTGGAAAGAAAGTATACATGAAATATATATACCACTATGATAAAG	116652
Qy	2584	AAAAATC-AAAAAATGTGTCAAGCCATTGTCAAGCTTGAATGGTCCCATGATCTACT	2641
Db	116653	AAAAACCAAAAAAATGTGTCAAGCAATGTCAAGCTTGAATGATCTGATGATCACT	116712
Qy	2642	TTTTCAATTGATATAAAGCCTCATATGATATGTTCACTGCTTATGTGATGCTTAG	2701
Db	116713	TTTTTCATTATTTAAAGCCTCATAAATGATGTTCACTTACTGATGTGATGCTTAG	116772
Qy	2702	CCCATTAATTGATTTTAAATCAGACAGCAATTACTTAACAGAACTTGAACAAGATGG	2761
Db	116773	CCCATTAATTGATTTTAAATCAGACAGCAATTACTTAACAGAAAGTGAACAACATGG	116832
Qy	2762	ACGTGATGAGAGGCTTAAATGTACTGATATGTGGAAGAGCCAAAGATCATCTCAG	2821
Db	116833	ACGTGATGAGAGGCTTATATGTATGTGATGTGGAAGAGCCAAAGATCATCTCAG	116892
Qy	2822	TAGATTTGTGTCAAAATACCTCTGAGCTCTGATTTGCCATGCTCTCATACAGAAA	2881
Db	116893	TGGATTTGTGTGTGGAACCTTCTGAGCTCTGATTTGTGCAAGTCCCATGACAGGAAA	116952
Qy	2882	TAAACAAGCTGTCCAGCATCTTGTGAACCTTGATTTGTCACACACTTTCAATTACATC	2941
Db	116953	TAAACAAGCTGTCCAGCATCTTCAATACCTGTAGTTGCTAACAGATTCACCTTACATC	117012
Qy	2942	CTGTAGGATCTCTCGAATTAAAGCAACAGAAAAGTCCCTGGAAGTCACTGAATCCAG	3001
Db	117013	CCGTAGGATCTCTCGAATTAAAGCAACAGAAAAGTCCCTGGAAGTCACTGAATCCAG	117072
Qy	3002	AAAGGCTCTTACCTTTAGCAACAGAGAGTCTTCAACCACTGACA----AAGAGGAAC	3057
Db	117073	AAAGGCTCTTACCTTTAGCAACAGAGAGTCTTCAACCACTGACA----AAGAGGAAC	117132
Qy	3058	GATTAAGGTAAGTACCAAGAACTCTTCTTCTTCCACATGATGATTTTGTCTGAAG	3117
Db	117133	GATTAAGGTAAGTACCAAGAACTCTTCTTCTTCTTAAATGATGATTTTGTCTGAAG	117192
Qy	3118	ATCATATCTTATATGCTTCCACCTTGATGTATACATGAGGAGTCCAGACTTGTTCAG	3177
Db	117193	ATCATATCTTGTGTCTTCCACCTTGATGTATACATGAGGAGTCCAGACTTGTTCAG	117252
Qy	3178	AAAAAGACAGAGACATGAAGCTTCTTTCAGAACTGATGTCTGTCAACCCAACTGTGT	3237
Db	117253	AAAAAGACAGAGACATGAAGCTTCTTTCAGAACTGATGTCTGTCAACCCAACTGTGT	117312
Qy	3238	GAGCTCTTAATGTGTCCCCCTTTTAATTATCTTCCCATATCACCTCTTCAATTCAA	3297
Db	117313	GAGCTCTTAATGTGTCCCCCTTTTAATTATCTTCCCATATCACCTCTTCAATTCAA	117372
Qy	3298	TCAATTCATCTGCCCTCATATGAGAGACATGCTGCCCTTACATTCATTTAAAGAGCAAGG	3357
Db	117373	TCAATTCATCTGCCCTCATATGAGAGACATGCTGCCCTTACATTCATTTAAAGAGCAAGG	117432
Qy	3358	GACATGACGAGATTTCTTCCAGAGTTGAACGTCTATAG-AGCCAGTTCTTGTGTTTAC	3416
Db	117433	GACATGACGAGCTTCTTCCAGAGTTGAACGTCTATAGAGCCAGTTTGTGCTTTTAC	117492
Qy	3417	TTACTTTCAATTATTTCTTCTTGTGCTATCTGAAAGTCTTAAAGAAATATGATGG	3476

Db 117493 TTACTTTCAAAATTTTCTCTCCATCTCGAATCGAAAGTTTCAAGGAAGACATGATGG 117552
 QY 3477 CCC-----AATATTTAAGAGTGTGTTTATGAGAAAGTATTTACAAAGATGACAGAG 3529
 Db 117553 TCCAAACTGAATGATGAGAGTGTTCATGAGGAGAGTATTTACAAAGACATGACAGAG 117612
 QY 3530 TTAA-----GGGTCAAGATCTTAAGCAGCAATACATAGGGGAGCACTA 3572
 Db 117613 TTAAAGAAAGAAATGATAGGTGTCAGACCTTAAGCAGCAATACCTTAGGGAGACATA 117672
 QY 3573 CTTCCTCCCTTAGAGGTGAAAGCAGACAGGAGAGAGATTTACCATTTGTC-----GCCAT 3626
 Db 117673 AGTCCTCCCTTAGATCTGAAGAGGAGCAGGAGAGAGATTTATTTATGTCACCATAGCCAT 117732
 QY 3627 AGCCATACCTGATGAGCATTAAGGTGGAGAGACATGACAGCAAGTGGAGAAAGCCCTGGG 3686
 Db 117733 AGCCATACCTGATCATAGGGGTGGAGAAACAAAGCGGAGAGGTGGAGAAAGCCCTGTA 117792
 QY 3687 TGGCCAAAGCAGACAGACAGAGCTGATATAGTTTGTGATCTGTGTCCCAACCAAAATCTC 3746
 Db 117793 TGGCCAAACACAGACAGACAGAGCTGATATGCTTTGATCTGTGTCTCCACCCAAATCTC 117852
 QY 3747 ATGTGATGTAATTTTCCAAATGTGAGAGAGGAGGAGGCTTGTGGAGATGATTTAGATCA 3806
 Db 117853 ATGTGATGTAATTTCCGATGTGGAGAGAGGAGGAGGCTTGTGGAGATGATTTAGATCA 117912
 QY 3807 CGGGATAGGTGATGATGATTTTAAACCAATCCCTCTGTGTGTGTGTGTGTGTGTGTGTGT 3866
 Db 117913 CGGGATAGGTGATGATGATTTTAAACCAATCCCTCTGTGTGTGTGTGTGTGTGTGTGTGT 117972
 QY 3867 CTGACGAGTCTCATGAAATCTAGTTGTTTAAAGAGGTGAGACCTCTCTC--CTCTCT 3924
 Db 117973 CTGATGATTTCTCATGATATCTAGTTGTTTAAAGGTGTAGACCTCTCTCTCTCTCTCT 118032
 QY 3925 CTTAATCTCTCTCTCAACATGTGAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3984
 Db 118033 CTTCCTCTCTCTCTCAACATGTGAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 118092
 QY 3985 GAAAGTTCTGAGGCT 4044
 Db 118093 GAAAGTTCTGAGGCT 118152
 QY 4045 GCTATTAGCCAGTAAACCATTTCTCTCATAAATTTCCAGTCTCAGATATTTCTTTT 4104
 Db 118153 GCGATGAACCATTTAAACCTGTTTCTCTATAAATTTCCAGTCTCAGATATTTCTTTT 118212
 QY 4105 AGCAATTTGAGATGAATTAATACACAGACAGAGGAGAGATGAAATCCCAAGTG 4164
 Db 118213 AGCAATTTGAGATGAATTAATACACAGAGGAGAGGAGAGATGAAATCCCAAGTG 118272
 QY 4165 CTTCCT 4224
 Db 118273 CT 118332
 QY 4225 CCAGAAATAAAGAAATCCCACTGATGTGTATCATAGAAAGCAGCTCTCTGGAGATGCA 4284
 Db 118333 CAAGAAATTAAGAAATCCCACTGATGTGTATCATAGAAAGCAGCTCTCTGGAGATGCA 118391
 QY 4285 ACAGATTAAGAAATGAGAAAGCAATCTCTGATGTGTATCATAGATCTCTCTCACT 4344
 Db 118392 ACAGATTAAGAAATGAGAAAGCAATCTCTGATGTGTATCATAGATCTCTCTCACT 118451
 QY 4345 TCTTCTATCT 4398
 Db 118452 TCTTCTATCT 118511
 QY 4399 TTTCACTCTATCAGACTATTTAATGTTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 4458
 Db 118512 TTTCACTCTATCAGAC-----TAAATGTTGAGATCTCTCTCTCTCTCTCTCTCTCTCT 118566
 QY 4459 CAGAAATGTTGATGTTTAAAGATGATAGAAATTAAGTAAATTTTAAAGAAAG-- 4515
 Db 118567 CAGATATGTTGCT 118626

QY 4516 -ACATATGATTTTGTGTTAAGATATAGTTGGCTGATCTATATAAGACATGAGAGAA 4574
 Db 118627 GACATATTTCTTTGTTGTTAAGATATAGTTGGCTGATCTATATAAGACATGAGAGAA 118686
 QY 4575 ATATCTTAAACAAGAAATAGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4634
 Db 118687 ATATCTT--AACAAGAAAGTACGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 118745
 QY 4635 TCAACAACTACAGAGAGAAATTTTCCAAAAATTAAGTGTGGGAGGTTGTATGTA 4694
 Db 118746 TCAACAAACACAGAGTGAAGAACTATTTTAAATTAAGTGTGGGAGGTTGTATGTA 118805
 QY 4695 CTGAACAGGTACAACTGTATTTCTTTGTCTCATATTTCTGAAAAATCACTAAT--ATAACA 4753
 Db 118806 CTGAACAGGTATTAACCTGTATTTCTTTGTCTCATATTTCTGAAAAATCACTAAT 118865
 QY 4754 AGAATCTATATAGCAATTTGCAATTTTGTCAATTTCTAATTAACCT--TAAATGATTTAAT 4812
 Db 118866 AGAATCTATATAGCAATTTGCAATTTTGTCAATTTCTAATTAACCTAATTTAAT 118925
 QY 4813 GTATCTGGAGAAAGTGCATAGATATATACAAATACCATAT-----ATPAGAAAT 4865
 Db 118926 GTATCTAGAGAAAGTGCATAGATATATACAAATACCATATGTTGTATTAAGAAAT 118985
 QY 4866 GAGCATCTGCAATTTTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 4925
 Db 118986 GAGCATCTGCAATTTTGT 119045
 QY 4926 AAAATGACACTTGTGAGATCTGAACTGAAAGCTCCAAAGCATCATATCATCAAGATTTCA 4985
 Db 119046 GAGATGACACTTGTGAGATCTGAAAGCTCCAAAGCATCATATCATCAAGATTTCC- 119104
 QY 4986 AAAATGCTGTCT 5045
 Db 119105 AAAATGAAAGCT 119164
 QY 5046 CCAGGACATTTAGCAATTTCCAGTCTTAAGAAAAAGAGAGGAGAGAGAGGCTTTGCT 5105
 Db 119165 CCAGGACATTTAGCAATTTCCAGTCTTAAGAAAAAGAGAGGAGAGAGGCTTTGCT 119223
 QY 5106 CCTTCTATTAATCCATGAG 5165
 Db 119224 TGTTCATTAATCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 119274
 QY 5166 ACAGCACTGTCTCATGAGATATCATACAGATCAAGAGAAATGGAGAGTGGGCTCTGT 5225
 Db 119275 AGAGCTCTGTCTCATATGATGCA--CCATCATCAAGAGAACTGTGAGTGGGCTCTGT 119333
 QY 5226 GCTGCTTGTACATTTCTCAAGAAAGTTATGTGACCAAAAAAGAGAAAT-----CTTGGAGCA 5280
 Db 119334 GCTGCTTGTACATTTCTCAAGAAAGTTATGTGACCAAAAAAGAGAAATTTATGGGAGCA 119393
 QY 5281 ACCAGAGTCTCTTCAAG 5340
 Db 119394 CTTAGAGAGTCTCTTCAAG 119453
 QY 5341 GGAATCTCTCAAGAAATTAATTAATGTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5400
 Db 119454 AGAATCTCTCAAGAAATTAATTAATGTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 119513
 QY 5401 CTTCAG 5459
 Db 119514 CTTCAG 119573
 QY 5460 TTTAATTTCTAG 5519
 Db 119574 TTTAATTTCTAG 119633
 QY 5520 TTTAAGAAAGATCATAGCATGTTTAAAGCAATGCTTAATGTTGGGATATCAAGAGTAA 5579
 Db 119634 TTTAA-----GAAAGCATGCTTAAAGTGTGATGATCTACAGTAA 119675

QY 5580 AAGCAGGACGACTTACTCTTACGATCTTCCGGTTTCATGAGAAACAATATC 5639
 DB 119676 AAGCAG-----CTTATGAGCTTCAGGTTTCATGAGAAATATATATC 119722
 QY 5640 ATACATACATATAGATGAGCAAAACAGTTAGTCTCTGAGTGTGATATACAGAGTTCT 5699
 DB 119723 ATACCCATATATATGAGCAAAAGTTTATGTCCTGATATGTGATATACAGATTTCT 119782
 QY 5700 CTTTTCTCTCCATTTCTTTTGGCCCATCAGAGCTGTGGCAGTTGTCTCCCTAAGA 5759
 DB 119783 CTTTTCTCTCCATTTCTTTGGCCCATCAGATGATATGAGAGCTGTCTCTACTA 119842
 QY 5760 GAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5819
 DB 119843 AAGGTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 119902
 QY 5820 CTTCTTTCACCTGAGAGCTCTGCGCATGATGATGATGATGATGATGATGATGATGAT 5879
 DB 119903 CTTCTTTCACATGAGAGCTCTGCGCATGATGATGATGATGATGATGATGATGATGAT 119962
 QY 5880 GCCTGTATGAGCAAAAGAACTCTGATTTTGTCTAATGAGCAGCATCTACCTCT 5939
 DB 119963 GCCTCAAGTATGAGCAAAAGAACTCTGATTTTGTCTAATGAGTTCATCTCTCT 120022
 QY 5940 AT-TTTTCCCTTATTTGATTTGCTTCTATCTCTATCTGAGATTTGTTAGTTAAT 5998
 DB 120023 ATATGTTTCCCTTATTTGATTTGCTTCTATCTCTATCTGAGATTTGTTAGTTA 120082
 QY 5999 TTTTTCACGCCCCCAATTTTGACTGCACTTGAATTTAATTGAGATGATGATGATGAT 6058
 DB 120083 TTTTTCACGCCCCCAATTTTGACTGCACTTGAATTTAATTGAGATGATGATGATGAT 120142
 QY 6059 CTTTACCCCTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6118
 DB 120143 CTTTACCCCTCTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120201
 QY 6119 GAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6178
 DB 120202 CAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120254
 QY 6179 CCCCCGACAAATATATCTCTTATATCTTATATTTCTATATTTCTGATATGAGAT 6238
 DB 120255 TCCGAGACAAATATATCTCTTATATCTTATATTTCTATATTTCTGATATGAGAT 120313
 QY 6239 AATATATATTTTATCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 6298
 DB 120314 AATATATATTTTATCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 120373
 QY 6299 TGTCAATCAGCTGTGTATTTTCCCAATTTAATGATGATGATGATGATGATGATGAT 6358
 DB 120374 TGTCAATCAGCTGTGTATTTTCCCAATTTAATGATGATGATGATGATGATGATGAT 120433
 QY 6359 TTGCATGATATCTCTCTTATATCTCTTATATTTCTATATTTCTGATATGAGAT 6418
 DB 120434 TTGCATGATATCTCTCTTATATCTCTTATATTTCTATATTTCTGATATGAGAT 120492
 QY 6419 CTCTCTCATACAAATGTTTGCATCAAAAGAACTCTACCTTCTCTATTTCTGATGAT 6478
 DB 120493 CTCTCTCATACAAATGTTTGCATCAAAAGAACTCTACCTTCTCTATTTCTGATGAT 120551
 QY 6479 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6538
 DB 120552 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120611
 QY 6539 ATTCCCATCTTCAATTAATCTGCTGAGTGCATCTGAGGCTCACTCTCACTTACTTTTC 6598
 DB 120612 ATTCCCATCTTCAATTAATCTGCTGAGTGCATCTGAGGCTCACTCTCACTTACTTTTC 120671
 QY 6599 CTCTGTTCTATGTTGAGCAATTCAGTCAAGTCCACACCTTCAACCAACATCTGTATC- 6657
 DB 120672 CTCTGTTCTATGTTGAGCAATTCAGTCAAGTCCACACCTTCAACCAACATCTGTATC- 120731
 QY 6658 AAAATGCTTACAGGGTTCCGCAAAAGCAGCATGATGCTTATTTTAAAGCAGATCATGAT 6717

DB 120732 AAAATGCTTACAGGGTTTCGCAAAAGCTACATGATGCTATGCTCAAAAGGCTCATGCT 120791
 QY 6718 GGTCAATTTTCAAGTTTGGGCACTCATCAATCAATTTCTCTCAACAGATATGAGTGTCC 6777
 DB 120792 GGTCAATTTTCAAGTTTGGGCACTCATCAATCAATTTCTCTCTCAACAGATATGAGTGTCC 120851
 QY 6778 ACAATATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6837
 DB 120852 ACAATATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120911
 QY 6838 CTTTACATGCTTCAATTTTATTTTCACTTTTCACTTCAATCAATGATGATGATGATGAT 6897
 DB 120912 CTTTACATGCTTCAATTTTATTTTCACTTTTCACTTCAATCAATGATGATGATGAT 120971
 QY 6898 CTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6957
 DB 120972 CTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121031
 QY 6958 ATGG-ATATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7016
 DB 121032 ATGGAATATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121091
 QY 7017 CACTGATGATTTTCTCAACCAAGCTCCGCGCATTTTCAAC-ATCTAGCGAAGATC 7074
 DB 121092 CACTGATGATTTTCTCAACCAAGCTCCGCGCATTTTCAACATTTATCTAGCAAAAGTTC 121151
 QY 7075 CCAATTTCTCTGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7134
 DB 121152 TCAATTTCTCTGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121211
 QY 7135 GTCATGAGGGTGTCAACAACCTCTTGTGTATGATGATGATGATGATGATGATGATGAT 7194
 DB 121212 GTCATGAGGGTGTCAACAACCTCTTGTGTATGATGATGATGATGATGATGATGATGAT 121270
 QY 7195 TTTCAGGCCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7254
 DB 121271 TTTCAGGCCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 121330
 QY 7255 TTTGATCTTATGTTATCTCCCATGTCAGACAGAACTTGTGTGAGATGATGATGATGAT 7314
 DB 121331 TTTGATCTTATGTTATCTCCCATGTCAGATGATGATGATGATGATGATGATGATGAT 121390
 QY 7315 GGTTCAGAGTCAACAAGAACTGATTTCAACTGATTTGAGAGACCCCACTTTTGTAT 7374
 DB 121391 GGTTCAGAGTCAACAAGAACTGATTTCAACTGATTTGAGAGACCCCACTTTTGTAT 121449
 QY 7375 AGTGAATTTATCTCTGAGTCTCTG-ATCTCTCTCTTT-AAATGAGACATTAATC 7432
 DB 121450 AGTGAATTTATCTCTGAGTCTCTG-ATCTCTCTCTTT-AAATGAGACATTAATC 121509
 QY 7433 CCACATGAGGAGGTGTGGGAGATCAGAGATCAACACAGTGTGATCACTGTGTTT 7492
 DB 121510 CCACATGAGGAGGTGTGGGAGATCAGAGATCAACACAGTGTGATCACTGTGTTT 121569
 QY 7493 CTGTTTCAGAGGTATCAAGCTGGGTTTCTGAGATGATTTCAACATCTCACTGTTGG 7552
 DB 121570 CTGTTTCAGAGGTATCAAGCTGGGTTTCTGAGATGATTTCAACATCTCACTGTTGG 121629
 QY 7553 GTACAGAACTGACACCAATCAAGAGAGTGAAGAGATCTCTGTTCAACAGACACCTGTA 7612
 DB 121630 GTACAGAACTGACACCAATCAAGAGAGTGAAGAGATCTCTGTTCAACAGACACCTGTA 121689
 QY 7613 GCTTCAAGGGGTGACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7672
 DB 121690 GCTTCAAGGGGTGACGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 121749
 QY 7673 TCTGCTCTGAGCTGTGCTGATGAGAGAGCTGTCTCATCTTCACTCACTCTG 7732
 DB 121750 TCTGCTCTGAGCTGTGCTGATGAGAGAGCTGTCTCATCTTCACTCACTCTG 121809
 QY 7733 TCCGCGCGCATCTCTCTTCTTATGCGGCAATATATGTTGCGCTTATGCTCATCA 7792

[illegible]

```

TITLE          The sequence of Homo sapiens clone
JOURNAL        Unpublished
REFERENCE       2 (bases 1 to 192618)
AUTHORS        Waterston, R.H.
TITLE          Direct Submission
JOURNAL        Submitted (04-JAN-2000) Genome Sequencing Center, Washington
                University School of Medicine, 4444 Forest Park Parkway, St. Louis,
                MO 63108, USA
COMMENT        MO 63108, USA
                On Jul 19, 2000 this sequence version replaced gi:2211362.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0206C01
----- Summary Statistics -----
Sequencing vector: M13, 93%
Sequencing vector: plasmid, 7%
Chemistry: Dye-terminator Big Dye, 7% of reads
Chemistry: Dye-terminator Big Dye, 7% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 179082 bases at least Q40
Consensus quality: 183120 bases at least Q30
Consensus quality: 185448 bases at least Q20
Insert size: 205000; agarose-fp
Insert size: 190318; sum-of-coverage
Quality coverage: 3.99 in Q20 bases; sum-of-coverage
Quality coverage: 4.34 in Q20 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 1280 1279: contig of 1279 bp in length
* 1380 1379: gap of unknown length
* 1380 3544: contig of 2165 bp in length
* 3545 3644: gap of unknown length
* 3645 5441: contig of 1797 bp in length
* 5442 5541: gap of unknown length
* 5542 8421: contig of 2880 bp in length
* 8422 8521: gap of unknown length
* 8522 10096: contig of 1575 bp in length
* 10097 10196: gap of unknown length
* 10197 11938: contig of 1742 bp in length
* 11939 12038: gap of unknown length
* 12039 14708: contig of 2670 bp in length
* 14709 14808: gap of unknown length
* 14809 17876: contig of 3068 bp in length
* 17877 17976: gap of unknown length
* 17977 21534: contig of 3558 bp in length
* 21535 21634: gap of unknown length
* 21635 24615: contig of 2981 bp in length
* 24616 24715: gap of unknown length
* 24716 28301: contig of 3586 bp in length
* 28302 28401: gap of unknown length
* 28402 35969: contig of 7568 bp in length
* 35970 36059: gap of unknown length
* 36070 42473: contig of 6404 bp in length
* 42474 42573: gap of unknown length
* 42574 49752: contig of 7179 bp in length
* 49753 49852: gap of unknown length
* 49853 57604: contig of 7752 bp in length
* 57605 57704: gap of unknown length
* 57705 67088: contig of 9384 bp in length
* 67089 67188: gap of unknown length
* 67189 77556: contig of 10468 bp in length
* 77557 77566: gap of unknown length
* 77567 92143: contig of 14387 bp in length

```

FEATURES

ORIGIN

Matches 5579; Conservative 0; Mismatches 725; Indels 125; Gaps 36;			
OY	1877	TGGGATAAAGAACACATTTTACACCTGCTGGTGGAAATGTAACTTGCGAACCACTATG	1936
DB	42354	TGTCAAAAAGAACGACTTACTACTCTGACAGCAAGATGTATATAGCACTACCTTGACG	42299
OY	1937	GAAAAAGTGTGGAAATTTCTTAAGGAAGCTAAAAAGTATGATGACCAATTGATCCAGCAAT	1996
DB	42294	CAAAACAATGGGAAAAATCTACATGAAATGAAAGTGAATCTACATTTTCATCATCAAT	42233
OY	1397	CCCAATTAATATGATATATATATAT-----ATT	2025
DB	42234	CCCAATTAAT	42177
OY	2026	TATATACCATGGAATACATCACTGACCCATATAAAAAAGATAAATATGATGACATTCACAGCA	2085
DB	42174	TATATACCATGGAATATATCTACAGCCATATAAAAAAGATATATATATATATATATATATAT	42111
OY	2086	CTTATGATGAAATTTGGAGACCCCTTATTTCTAAAGTGGGTAACTCAGAAATGGAAAAACAAC	2145
DB	42114	CCATGATGGAACCTGGACA-TATTACTCTAACCTGAGGTAACTCAGAAATGGAAAAACAAC	4205
OY	2146	ATCATATGTTCTCACTTACAAAGTGGGGCTTAACCTGTAGAGACACCAAGCATATGAATGA	2205
DB	42055	ATCATATATATTTCTCACTTACAAAGTGAAGGCTTAACTGTGAAGATGTAAATCATATGAATAA	4199
OY	2206	TATATATGAACCTGTGGGACCTTGAGGGGAAGATGGAAGAAGGCGAGGATATAAAGACTA	2265
DB	41995	TATATATGAACCTATGGGGACCTTCAGGGGAAGATGGAAGAAGGGTGAGGGATATAAAGACTA	41933
OY	2266	CACAAATGGGTACAGTGTACACTGCTCAGTGTATGGGTGACCAAAATCTCAGAAATTTACC	2325
DB	41935	CACATTTGGGTATAGTGTACACTGCTCAGTGTATGGGTGACCAAAATCTCAGAAATTTACC	41877
OY	2326	ACTTAAAGACCTTATCCATGGAAGAAACACCACTGTTCCTCCCAAAATCCCAATGAATAA	2385
DB	41875	ACTTAAAGACCTTATCCATGGAAGAAACACCACTGTTCCTCCCAAAACCCATTTGAATAA	4181
OY	2386	AAAT	2444
DB	41815	AAAT	41755
OY	2445	TTATATATAT-----GAAGTAAAGATGAAAAGTGTAGCAAGTGTAGCCTCTGGACAATGG	2500
DB	41755	TGAT	41699
OY	2501	GACTCTAGATTTTCACTTGCAATACACTTCTGTGGCACTTTGGAAGAAGATATACATG	2560
DB	41695	CATCTCTAGATTTTCACTTGCAATACACTTCTCTACCACTTGGAAAAGAACATATACATG	41633
OY	2561	AATATATACACAT	2620
DB	41635	AATATATACACAT	41577
OY	2621	TGAATGTCTCCATGATCTACTTTTCTATTGATATATAAAGCCTCATATATGATAGTTTACA	2680
DB	41575	TG-ATGATCTCTATGATCTACTTATTTCTTATAGTTAAABAGCCTCATATCAAAAGTTTACA	41511
OY	2681	TTGCTTATATGATGCTTGAAGCCCATATATGATTTTAAATCAGACACAGCAATATCTTA	2740
DB	41516	TTACTTATATGATGATATCTTGAAGCCCATATATGATTTTCAAAATCAGACACAGCAATATCTTA	4145
OY	2741	CAGGAATTTGAAACAAAGTGGGAAGCTATATAGAGAGCTTAAATGTAATCTGATATGGACA	2800
DB	41456	CAGGAATCTGAAACAAAGTGGGAAGCTATATAGAGAGCTGTAATGTAATGTAATGGACA	4139
OY	2801	GAGGCAAGATATCTCACTTATGAAATTTGTCTCAAAATACCTCTGGCCTCTGATTTGC	2860
DB	41396	GAGGCAAGATATCTCACTTATGAAATTTGTCTCAAAATACCTCTGGCCTCTGATTTGT	4133
OY	2861	CCATATGTTCTCATACAGAAATTAACAAGCTGTCCAGACATCTTTCGTAAAGCTGATTTGCT	2920
DB	41336	CCATATGTTCTCATACAGAAATTAACAAGCTGTCTCAGACATCTTTCATTAAGCTGATTTGCT	41277

QY 2921 CACCAGCTTTGATTTCAAGCTCCGTAGGATCTCTGAATTAGCAACAGAAAAGTCC 2980
 DB 41276 AACCGCTTTCAATTTCAAGCTCTGTAGGATCTCTGAATTAGCAACAGAAAAGTCC 41217
 QY 2981 TCTGAAGTCACTGAATCCAGAAAAGGCTCTACCTTTAGCAACAGGAGGTCTTACCA 3040
 DB 41216 TCTGAAGACTGAATCCAGAAAAGGCTCTACCTTTAGCAACAGGAGGTCTTACCA 41157
 QY 3041 CTGAGC-AAAGAAAGGACATAGGAGTAAAGTACCAAGAACTCTTCTTCCAGATCACT 3099
 DB 41156 CAGGCAAAAGAAAGAAATAGGAGTAAAGTACCAAGAACTCTTCTTCCAGATCACT 41097
 QY 3100 TATGATTTTGGCTGTAAAGATCAATGTCTTATGCTTCCACCTTGAGTCAATGACAGGAG 3159
 DB 41096 TATGATTTTGGCTGTAAAGATCAATGTCTTCCACCTTGAGTCAATGACAGGAG 41038
 QY 3160 TCACGAGCTTGTTCAGAAAAGACAGAGACATGAAGCTTCTTTCAGAACTGAGTGC 3219
 DB 41037 TCATGAGCTTGTTCAGAAAAGACAGAGACATGAAGCTTCTTTCAGAACTGAGTGC 40978
 QY 3220 TGTCAACCCAACTGTGTAGCTCTAAATGTGTCCCTTCTTAATTTATCTCCCAT 3279
 DB 40977 TGCCAGCCCAACTGTGTAGCTCTAAATGTGTCCCTTCTTAATTTATCTCCCAT 40919
 QY 3280 TCACCTCCCTTCATCCAACTTCATTCATCTGCCCTCAATGAGAGACGTGCTTACAT 3339
 DB 40918 TCATCTCTCTTCATCCAACTTCATTCATCTGCCCTCAATGAGAGACGTGCTTACAT 40859
 QY 3340 TCATTTAAAGAGCAAGGAGACATGACAGCAATTTCTTCCAGAGTTAACTGCTAATAGC 3399
 DB 40858 TCATTTAAAGAGCCAGAAACATGAAGACATTTCTTCCAGAGTTAACTGCTAATAGC 40799
 QY 3400 CAGTTTCTTGTTCACCTTACTTCTTCAATTTATCTTCTTGTCCCTATCTGAAAAGTCT 3459
 DB 40798 CGGTTTATGTGTTCACCTTACTTCTTCAATTTATTTCTTGTCCCTATCTGAAAAGTCT 40739
 QY 3460 AAGAAAGATATAGATGAGCCAA-----TAATTAAGAGTGTTCATGAGAAAGTAT 3512
 DB 40738 AAGAAAGTATATAGATGATCCAAATGAATTAATTAAGAGTGTTCATGAGAAAGTAT 40679
 QY 3513 TACAAAGATGACAGAGTTAA-----GGGTCAAGATCTTAAGACGA 3554
 DB 40678 TACAAAGATGACAGAGTTAAAGGAAAGATGAGGATGAGAAACGACTCAAGACGA 40619
 QY 3555 ATACATAGGAGGACCTACTTCTCTCCCTAGGCTGAAGAGGAAAGGAAAGACATATAC 3614
 DB 40618 ATACATAGGAGGACCTACTTCTCTCCCTAGGCTGAAGAGGACAGGAAAGACATATAC 40559
 QY 3615 CATTGTCCCATAGCACTAGCTGTAGCCATTAAGGAGTGGAGACATGAGGACAGGAGTGG 3674
 DB 40558 CATTGTCCCATAGCACTAGCTGTAGCCATTAAGGAGTGGAGAGACAGGACAGGAGTGG 40499
 QY 3675 AGAAGCCCTGCGTGGCCAAAGCAAGCAACAGGCTGATATATGATGATCTGTGTCC 3734
 DB 40498 AGAAGCCCTGCGACAGCAACCAACAGCAACAGGCTGATATGATGATCTGTGTCC 40439
 QY 3735 CACCAAAATCTCATGTGATTTGTAATTTCCAAATGTTGAGAGGAGGCTTGTGGAGTGG 3794
 DB 40438 TACCAAAATCTCATGTGATTTGTAATTTCCAAATGTTGAGAGGAGGCTTGTGGAGTGG 40381
 QY 3795 ATTATTAGATCAGGAGATGATTTTGCATGAATGTTTAAACACATCCCTTTGGAT 3854
 DB 40380 ATTATTAGATCAGGAGATGATTTTGCATGAATGTTTAAACACATCCCTTTGGAT 40321
 QY 3855 GTTGTGTGATATCTGACAGTTCATGAATCTAGTTGTTTAAAGGCTGTGACATCTC 3914
 DB 40320 ATTGTGTGATATCTGACATGATTTCTCATGAATCTAATGTTTAAAGTGTGACATCTC 40261
 QY 3915 CTCCT 3972
 DB 40260 CTCCT 40201

QY 3973 TCACAGAGATGGAAGCTCTGAGGCTCCCAAGACAGAGAGCTATGCTTCT 4032
 DB 40200 TCACAGAGATGGAAGCTCTTGAAGCTCCCTTAAGAGAGAGCTATGCTTCTCT 40141
 QY 4033 ACAGTGTGAGACTATTAAGCAAGTTAAACCATTTCTTCAATTAATTTCCAGTCTGAG 4092
 DB 40140 ATAGCCTGTAGAGCCATGAGCAATTAACCTGTTTCTCTCAATTAATTTCCAGTCTGAG 40081
 QY 4093 GTATTTCTTTTGAATTTGAAATGAACTAATACAGACAGAGAGCCAGAGATGGA 4152
 DB 40080 GTATTTCTTTTGAATTTGAAATGAACTAATACAGAGAGAGCCAGAGATGGA 40021
 QY 4153 AATCCAGAGTCTTCTCTGCTCTTCCAGCTCTCTGCTGTCTCCAGTGTCTCA 4212
 DB 40020 AATCCAGAGTCTCTCTGCTCTTCCAGCTCTCTGCTGTCTCCAGTGTCTCA 39961
 QY 4213 TTCACAGAAAACAGAAAATAAAGAAATCCACTGATGTGTATCAAGAACCACTCTC 4272
 DB 39960 TTTCAATTAGAAAACAGAAAATAAAGAAATCTAATGTGTATCAAGAACCACTCTC 39901
 QY 4273 TTGGAGTGTCAACAGGATTAAGAAAGAAAGCAAAATCTCATGTTAAATGAGACTA 4332
 DB 39900 TTGGAGTGTCAACAGGATTAAGAAAGAAAGCAAAATCTCATGTTAAATGAGACTA 39841
 QY 4333 TCCCTCTCACTTCTGTATCTCTCTAATTCCTGGGGCTTCTCT-----ATCTGATTG 4386
 DB 39840 TTTCTCTTACCTTCTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 39781
 QY 4387 ATCCGTGTCAATTCAGCTCTATCACTACTTAAATGTTGCTGTCTTCTCTACT 4446
 DB 39780 ATTCCTGTGTCAATTCAGCTCTATCACTACTTAAATGTTGCTGTCTTCTCTACT 39721
 QY 4447 GTCACTTTATGAGAAATGTTGATTTGTTAAATAAGCATAGAAATTAATATGTAAT 4506
 DB 39720 GTCACTTTATGAGAAATGTTGATTTGTTAAATAAGCATAGAAATTAATATGTAAT 39661
 QY 4507 TTAATAAGAAATATGATTTTGTGTAGAAATAAGTTGGCTGATCTTAATAAGACATG 4566
 DB 39660 TTAATAAGAAATATGATTTTGTGTAGAAATAAGTTGGCTGATCTTAATAAGACATG 39601
 QY 4567 AAGAAAGAAATCTTAAACAGAAAGTATGTTGCTGTGGTCACTAGGTTCTGAAT 4626
 DB 39600 AAGTAAATATCTTAAACAGAAAGTATGTTGCTGTGGTCACTAGGTTCTGAAT 39541
 QY 4627 CTACAGATTCACAACTACAGAGAGAACTTTCCAAAATAAAGTGTGGCGAGTTG 4686
 DB 39540 CTGCAATTCACAACTACAGAGAGAACTTTCCAAAATAAAGTGTGGCGAGTTG 39481
 QY 4687 TGTATGTACTGAACAGGTACAACTGTATTTCTTGTCTATTTCTGAAAACTACAA 4746
 DB 39480 TGTATGTACTGAACAGGTACAACTGTATTTCTTGTCTATTTCTGAAAACTACAA 39421
 QY 4747 TATACAGAACTTATATAGCAATTTGCAATTTGTGCAATTTCTAATACTTAAATGA 4806
 DB 39420 TATACAGAACTTATATAGCAATTTGCAATTTGTGCAATTTCTAATACTTAAATGA 39361
 QY 4807 TTTAATGTATCTGGAGAAAGCAATAGCAATATACAAATACATATTAAGAAATG 4866
 DB 39360 TTTAATGTATCTGGAGAAAGCAATAGCAATATACAAATACATATTAAGAAATG 39302
 QY 4867 AGCATCTGAGATTTTGTCTGTGCTGGGTTCTGAAAAGAAATCCCTGTAAATACAA 4926
 DB 39301 AGCATCTGAGATTTTGT 39242
 QY 4927 AATGACACTCTTGCAGATCTGAACAGAGCTCCAAAGCATATACATCAAGATTCGA 4986
 DB 39241 AATGACACTCTTGCAGATCTGAACAGAGCTCCAAAGCATATACATCAAGATTCGA 39182
 QY 4987 AATGTGCTGCTCCCACTCTGAGAGTGTGCTCATCTCTGTGATCTTCAATGTTCC 5046
 DB 39181 AATGTGCTGCTCCCACTCTGAGAGTGTGCTCATCTCTGTGATCTTCAATGTTCC 39123
 QY 5047 CAGCAGATTAACATTCAGTCTTATGAAAAAGAACAGAGGAGAGAGCTTGTCTC 5106

Db 39122 CAGCTACATGAGATTCACGCTTATGG-AAAAGAGACAGGGGAGAGGCTTTATCTC 39064
 Qy 5107 CTCTTAATTAATCCATAGGACAGGACTTGCTGTGCACTTTTGATTTCTTCCACTTAA 5166
 Db 39063 CTCTTAATTAATCCATAGGACAGGAAATTGCTTACATCAC-TTGTGATTTCTTCCACTTAA 39005
 Qy 5167 CAGACCTGCTCATGGAGATTCATCCAGCATCAAGAAACCTGGATGTGGGCTTTGTG 5226
 Db 39004 CAGACCTGCTCATCAATGTCATCCAGCATCAAGAAACCTGGATGTGGGCTTTGTG 38945
 Qy 5227 CTGCTGTACATCTCGAAGGTTATGTACCA-----AAAAGAAATCTTGGGGCAA 5281
 Db 38944 CTGCTTTTACTCTCGAAGTTATGTACCGAAGAGAAATGAAATTTGGGGCAAT 38885
 Qy 5282 CCAGAGCTCTTCCAGCCCTGACGTCTGTGATTTCTGTGCTCATCAAGATTTTTCAG 5341
 Db 38884 CTGAGAGTCTTTCAG-CCCTGATGTCTGTGACTGTGCTCATCAAGATTTTTCAG 38826
 Qy 5342 GAACTCTCGAATTAATTAATGTGGGGCAGAGACAGAACTGGAGTCTGTGCGAGAC 5401
 Db 38825 GAACTCTCGAATTAATTAATGTGGGGCAGAGACAGAACTGGAGTCTGTGCGAGAC 38766
 Qy 5402 TCAGAGGACACAGGGGCTGTATTTGACCTGTCTCATGTTTGTGAAC-CAGGAAACCT 5460
 Db 38765 TCAGAGGACACAGGGGCTGTATTTGACCTGTCTCATGTTTGTGAACCTCAGGAAACCT 38706
 Qy 5461 TTAATTTCTAGGACCTTAGCTTCATTTATGTTATATAGAGATTAATACATAGACGT 5520
 Db 38705 TTAATTTCTAGGACCTTAGCTTCATTTATGTTATATAGAGATTAATACATAGACGT 38647
 Qy 5521 TTAAGAACTCATAGCATTTAAACAACTGCTAAATGTTGTGATACCAAGTGAATA 5580
 Db 38646 TTAAGAACTCATAGCATTTAAACAACTGCTAAATGTTGTGATACCAAGTGAATA 38605
 Qy 5581 AGACAGGACATGACTTACTCTTAACGATCTTCCGGGTTTCAATGAGGAAAGAAATATCA 5640
 Db 38604 TGAAGGATGACTTACTCTTAATGAGCTTGGTTTCAATGAGGAAAGAAATATCA 38545
 Qy 5641 TACCATACCTATAGATGAGACAAACAGTTAGTGTCTGATGTGATTAACAGAGTTCTC 5700
 Db 38544 TACCATACCTATAGATGAGACAACTTTAGTGTGTGCTGATGTGATTAACAGAGTTCTC 38485
 Qy 5701 CTTTTCTCTCCATTTCTTTTGGGCAATCAAGCTGTGGCAGCTTGTCTCCCTAAGAG 5760
 Db 38484 CTTTTCTCTCCATTTCTTTTGGGCAATCAAGCTGTGGCAGCTTGTCTCCCTAAGAG 38425
 Qy 5761 AGCTCATGATGATGACATCACTCTGATGATGCTCTCTATATCTCCAGAGAGATGATC 5820
 Db 38424 AGCTCATGATGATGACATCACTCTGATGATGCTCTCTATATCTCCAGAGAGATGATC 38365
 Qy 5821 TTTCTTTCACCTGAGAGCTCTGCCATGTGCAATTTTGGGATTCAGAGACAAACGTGG 5880
 Db 38364 TTTCTTTCACCTGAGAGCTCTGCCATGTGCAATTTTGGGATTCAGAGACAAACGTGG 38305
 Qy 5881 CCTCTGATAGGCAAAAAAGAACTCTGATTTGTTCTTAATGAGCAGCATCACTCTTA 5940
 Db 38304 CCTCTGATAGGCAAAAAAGAACTCTGATTTTCTTAATGAGCAGCATCACTCACTCTTA 38245
 Qy 5941 TTTT-TCCCTTATTTCAATTTGCTTCTCATTTCTGATGTGAGATTTTGTAGGTTAATTT 5999
 Db 38244 TTTTGTCCCTCTTTGATTTGCTCTCTCTCTCATATGTGAGATTTGTAGG-TCAAGTTT 38186
 Qy 6000 TTTTTCAGCCCAAAATTTGACTGTGCACTTGAATTTAACTTGAGATCACTCTCTTAC 6059
 Db 38185 TTTTTCAGCCCAAAATTTTGGGCTGCAATTTGGAATTTAACTTGAGATCACTCTCTG 38126
 Qy 6060 TTTTACCCCTCTTAACATGTATATGACACATAGTGTGCTGGGCTCAAAAGGGCTGTG 6119
 Db 38125 TTTT-TTCCCTGTAACTATATATGACACAGGCTGTGCTGGGTTAAAGGGCTGTGCTC 38067
 Qy 6120 AAAAAATGATCATGATGACCTGTGGCTCACTTCAATTAATATATATATACC 6179

Db 38066 CAAAGCTGATTTGTGACTCAGACCTGCGGACCTCAAAATTCATATATATATATACC 38007
 Qy 6180 CCCCAGCAAAATATATATCTCTCTTTTATCTTAATTTCAATTAATCTCAATATACAGATA 6239
 Db 38006 CTTGGGCAAAATATATATCTCTCTTTTATCTTATATATATATATATATATATATAT 37947
 Qy 6240 ATACTAATTTTATCTCTCTGAGCTCTTCAATGATTTAAAGAGCAATATCTAATTAAT 6299
 Db 37946 ATACTAATTTTATCTCTCTGAGCTCTTCAATGATTTAAAGAGCAAACTTCAAACT 37887
 Qy 6300 GTCAATGAGCTGTATATCTCCAAATTAAGCTTAATCTCATCTTCCAGTTGAATTT 6359
 Db 37886 GTCAATGAGCTGTATATCTCCAAATTAAGCTTAATCTCATCTTCCAGTTGAATTT 37827
 Qy 6360 TGCATGATATCTCTCTTTTCAACCCAGGCTTACATCTTCTCTTATTTTCACTATGAGAC 6419
 Db 37826 TGCATGATATCTCTCTTTTCAACCCAGGCTTACATCTTCTCTTATTTTCCCTCATGAGAC 37767
 Qy 6420 TCTCTCATACAAATTTTGCATCAACAAAGAAAGCTTACCAAGATCTCCGAAAGAGA 6479
 Db 37766 TCTCTCATACAAATTTTGCATCAACAAAGAAAGCTTACCAAGATCTCCGAAAGAGA 37707
 Qy 6480 GAATGAATATGTTTACATTTGTATATCTGAGAGAACTTTAGTATCTCCCATATATA 6539
 Db 37706 GAATGAATATGTTTACATTTGTATATCTGAGAGAACTTTAGTATCTCCCATATATA 37647
 Qy 6540 TTCCCACTTCAATTAATCTGTGCTGAGTGGCACTGAC-GETCACCTCTCACTTACTTCTTC 6598
 Db 37646 TTTCACACTTCAATTAATCTGTGCTGAGTGGCACTGAGTGGCACTTCACTTACTTCTTC 37587
 Qy 6599 CTCTGTTCTATTTGCTGAGCAATTCAGTCAAGCCACACCTTACCCAAACATCTGTATC- 6657
 Db 37586 CTCTGTTCTATTTGCTGAGCAATTCAGTCAAGCCACACCTTACCCAAACATCTGTATC- 37527
 Qy 6658 AAAATGCTTACAGGGGTTGGGCAAGCCACATGAGCTTATTTTAAAGGCAATAGT 6717
 Db 37526 AAAATGCTTACAGGGGTTGGGCAAGCCACATGAGCTTATTTTAAAGGCAATAGT 37467
 Qy 6718 GGTCAATTTACAGGTTTGGGCACTCAATCAATCTTCTTCAACACAGATAGAGCTTCC 6777
 Db 37466 GGTCAATTTACAGGTTTGGGCACTCAATCAATCTTCTTCAACACAGATAGAGCTTCC 37407
 Qy 6778 ACAATATGATTTGATGATTAATTTTCTTCAATTAATATATATATATATATATAT 6837
 Db 37406 ACAATATGATTTGATGATTAATTTTCTTCAATTAATATATATATATATATATAT 37347
 Qy 6838 CCTTACATTTGCTTCAATTTTATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTT 6897
 Db 37346 CCTTACATTTGCTTCAATTTTATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTT 37288
 Qy 6898 CTTCTCATGATGATTTCTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTT 6957
 Db 37287 -TCTCATGATGATTTCTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTT 37230
 Qy 6958 ATGG-ATATGAGCACTAGAGATTAATTAATGATGATGATGATGATGATGATGATGAT 7016
 Db 37229 ATGGAATATGAGCACTAGAGATTAATTAATGATGATGATGATGATGATGATGATGAT 37170
 Qy 7017 CACTGATATTTTCTCAACCCAGAGGTCGCAATTTTCAAC-ATCTGAGGAAGATC 7074
 Db 37169 CACTGATATTTTCTCAACCCAGAGGTCGCAATTTTCAACCTTATCTTCAAGAGGTC 37110
 Qy 7075 CCAATTTCTCTGCTTCAATATGATGATTTTCTCTGCTGAGAGATGACAGTCTG 7134
 Db 37109 ACAATTTCTCTGCTTCAATATGATGATTTTCTCTGCTGAGAGATGACAGTCTG 37050
 Qy 7135 GTCATAGGATGATCAACCACTCTTTTGTATATGATTTTCTCTCACTGAGAGAAAT 7194
 Db 37049 GTCATAGGATGATCAACCACTCTTTTGTATATGATTTTCTCTCACTGAGAGAAAT 36990
 Qy 7195 TTCAGGCCAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7254
 Db 36989 TTCAGGCCAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 36930


```
* 36070 /42473: contig of 6404 bp in length
* 42474 /42573: gap of unknown length
* 42574 /49752: contig of 7179 bp in length
* 49753 /49852: gap of unknown length
* 49853 /57604: contig of 7752 bp in length
* 57605 /57704: gap of unknown length
* 57705 /67088: contig of 9384 bp in length
* 67089 /67188: gap of unknown length
* 67189 /77656: contig of 10468 bp in length
* 77657 /77757: gap of unknown length
* 77757 /92143: contig of 14387 bp in length
* 92144 /92243: gap of unknown length
* 92243 /109240: contig of 16997 bp in length
* 109241 /109340: gap of unknown length
* 109341 /124285: contig of 14945 bp in length
* 124286 /124385: gap of unknown length
* 124386 /138608: contig of 14223 bp in length
* 138609 /138708: gap of unknown length
* 138709 /153882: contig of 15174 bp in length
* 153883 /153982: gap of unknown length
* 153983 /169994: contig of 16012 bp in length
* 169995 /170094: gap of unknown length
* 170095 /192618: contig of 22524 bp in length.
```

FEATURES

```
source
1..192618
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP11-206c1"
1..1279
/note="assembly_name:Contig19"
1380..3544
/note="assembly_name:Contig21"
3645..5441
/note="assembly_name:Contig22"
5542..8421
/note="assembly_name:Contig23"
8522..10096
/note="assembly_name:Contig24"
10197..11938
/note="assembly_name:Contig25"
12039..14708
/note="assembly_name:Contig26"
14809..17876
/note="assembly_name:Contig27"
17977..21534
/note="assembly_name:Contig28"
21635..24615
/note="assembly_name:Contig29"
24716..28301
/note="assembly_name:Contig30"
28402..35969
/note="assembly_name:Contig31
clone_end:SP6
vector_side:left"
36070..42473
/note="assembly_name:Contig32"
42574..49752
/note="assembly_name:Contig33"
49853..57604
/note="assembly_name:Contig34"
57705..67088
/note="assembly_name:Contig35"
67189..77656
/note="assembly_name:Contig36"
77757..92143
/note="assembly_name:Contig37"
92244..109240
/note="assembly_name:Contig38"
109341..124285
/note="assembly_name:Contig39"
124386..138608
```

```
misc_feature /note="assembly_name:Contig40"
138709..153882
/note="assembly_name:Contig41"
misc_feature /note="assembly_name:Contig42"
153983..169994
/note="assembly_name:Contig43"
170095..192618
/note="assembly_name:Contig43"
ORIGIN
Query Match 52.1%; Score 4492.2; DB 2; Length 192618;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 4555; Conservative 0; Mismatches 13; Indels 5; Gaps 5;
1 TGTATGAGCCCATGTCCTTAAATATACCAAAACCGAGAAAGATATCAAAAAAGAAAC 60
53037 TGTATGAGCCCATGTCCTTAAATATACCAAAACCGAGAAAGATATCAAAAAAGAAAC 53096
61 TATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
53097 TATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53156
121 CCAATCCAGACGATATCAAGATATCAAGATATCAAGATATCAAGATATCAAGATATCA 180
53157 TCAATCCAGACGATATCAAGATATCAAGATATCAAGATATCAAGATATCAAGATATCA 53216
181 GTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
53217 GTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53275
241 AAAACAAATATCATATATCATATATCATATATCATATATCATATATCATATATCATAT 300
53276 AAAACAAATATCATATATCATATATCATATATCATATATCATATATCATATATCATAT 53335
301 ATTTCTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 360
53336 ATTTCTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 53395
361 TAAAGCCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 420
53396 TAAAGCCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 53455
421 CATTCCTCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480
53456 CATTCCTCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 53515
481 AGTAGGAAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
53516 AGTAGGAAAGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 53575
541 CAATTAAGAGAACTCAAACTGTCCTGTTCACTGATGATGATGATGATGATGATGATGATG 600
53576 CAATTAAGAGAACTCAAACTGTCCTGTTCACTGATGATGATGATGATGATGATGATGATG 53635
601 CCTTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
53636 CCTTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 53695
661 ATACAACTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
53696 ATACAACTAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 53755
721 GAATCAATCAAGAACTCAAACTGTTTAAATGCTGTAATAAATACTTAAGAAATTT 780
53756 GAATCAATCAAGAACTCAAACTGTTTAAATGCTGTAATAAATACTTAAGAAATTT 53815
781 CTTACCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
53816 CTTACCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 53875
841 GATGACCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
53876 GATGACCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 53935
```

QY	901	AAAAAGCATTATTTGCCAAAAGCATCTTCAAGTTCAATGCATATCCACCAAAATATCA	960
Db	53936	ATATATGCCATATATTTGCCAAAAGCAATCTACAAAGTTCAATGCATATCCACCAAAATATCA	539959
QY	961	TCATCATTTCTCCACAGAACTTAGAAAAAAACAATCTTAATAATTCAATGTGAACAAACCA	1020
Db	53996	TCATCATTTCTCCACAGAACTTAGAAAAAAACAATCTTAATAATTCAATGTGAACAAACCA	540555
QY	1021	AAAAAAAAAAAAAAAAAACCAGATAGCCAAAGCAAGACTTAGCAAAAAGAACAAATCTGGA	1080
Db	54056	AAAAAAAAAAAAAAAAAACCAGATAGCCAAAGCAAGACTTAGCAAAAAGAACAAATCTGGA	541155
QY	1081	GGCATCATTTACCATTCCTTCAAACTATATCTAACAGGCTTATATACCAAAAATCATATGG	1140
Db	54116	GGCATCATTTACCATTCCTTCAAACTATATCTAACAGGCTTATATACCAAAAATCATATGG	541757
QY	1141	CATGTGATTAACCTAGAGCATATGACCAATGGAATAAAGAGAAATCTCAAAATTTAAG	1200
Db	54176	CATGTGATTAACCTAGAGCATATGACCAATGGAATAAAGAGAAATCTCAAAATTTAAG	542359
QY	1201	CCAAATTAATTTATAGCCCAATGATTTTGGACAAAGCAAAACAAATAAAGTGGGGAAAA	1260
Db	54236	CCAAATTAATTTATAGCCCAATGATTTTGGACAAAGCAAAACAAATAAAGTGGGGAAAA	542959
QY	1261	GACATCTTAGTTAACAAAATGCTGAGATTAATGGCAAGCCACATGTGGAAATGAATGA	1320
Db	54296	GACATCTTAGTTAACAAAATGCTGAGATTAATGGCAAGCCACATGTGGAAATGAATGA	543555
QY	1321	CTGGATCCCTGTGCTCCTTAATATACAAAATTTGAATACAAAGATGGATCAAAAGCTTAA	1380
Db	54356	CTGGATCCCTGTGCTCCTTAATATACAAAATTTGAATACAAAGATGGATCAAAAGCTTAA	544144
QY	1381	TCTGAGACCTAAAAACCAATAAAATTTCTGAAGATTAATCAAGAAAAATGCTTACAGAT	1440
Db	54415	TCTGAGACCTAAAAACCAATAAAATTTCTGAAGATTAATCAAGAAAAATGCTTACAGAT	544744
QY	1441	TCATTTAGCCAAAGACTTCATGCGCCAGAACCCCAAAGTAAATGCACAAAAACAAAAAT	1500
Db	54475	TCATTTAGCCAAAGACTTCATGCGCCAGAACCCCAAAGTAAATGCACAAAAACAAAAAT	545344
QY	1501	AAATGATTAAGACTTAATTAATCTAAAAAGCTTTGGCAGAGAAAACAATATTAAGAG	1560
Db	54535	AAATGATTAAGACTTAATTAATCTAAAAAGCTTTGGCAGAGAAAACAATATTAAGAG	545939
QY	1561	AGCAAAACAGAAACCCACCGAGTGGAGAAAAATCTTCAAACTAAGCATCTGACTTAAGG	1620
Db	54594	AGCAAAACAGAAACCCACCGAGTGGAGAAAAATCTTCAAACTAAGCATCTGACTTAAGG	546555
QY	1621	ACTAATATTCGGAAATCCAAAGGAATCTCAAAACAAATCAGCAAGAAAGAAAGCAACATCC	1680
Db	54654	ACTAATATTCGGAAATCCAAAGGAATCTCAAAACAAATCAGCAAGAAAGAAAGCAACATCC	547133
QY	1681	CATGAAGAATGGCTTAAGGACATGAATAGCAATTTCTCAAAAGAAAGATATCAAAATGGC	1740
Db	54714	CATGAAGAATGGCTTAAGGACATGAATAGCAATTTCTCAAAAGAAAGATATCAAAATGGC	547737
QY	1741	CAACAAACAGAAAAAAATGCTTAAATCATCTTAATGATTAAGGAAATGTAAATCAACATGG	1800
Db	54774	CAACAAACAGAAAAAAATGCTTAAATCATCTTAATGATTAAGGAAATGTAAATCAACATGG	548333
QY	1801	TAAATGCAATACACCTTATCTCTGCAAGAAATGTCATATTTAAAAATCTAAAAATTAATA	1860
Db	54834	TAAATGCAATACACCTTATCTCTGCAAGAAATGTCATATTTAAAAATCTAAAAATTAATA	548933
QY	1861	GATGTGTGTGGTCTGTGTGTATTAAGGAACATTTTACATCTGTGTGTGGAAATGTAAAC	1920
Db	54894	GATGTGTGTGTGTGTGTGTATTAAGGAACATTTTACATCTGTGTGTGGAAATGTAAAC	549533
QY	1921	TTGCGCAACACATTTGGAAAAACAGTGTGGAAATTTCTTAAAGAACTAAAGATAGATCGAC	1980
Db	54954	TTGCGCAACACATTTGGAAAAACAGTGTGGAAATTTCTTAAAGAACTAAAGATAGATCGAC	550133
QY	1981	CATTGTATCCAGCAATCCCATTAATATATATATATATATATATTTATATACATGGAAT	2040

Db	55014	CATTGATCCAGCAATCCCAATTAATGTATTAATATATATTTATATACATGAAT	55073
OY	2041	ACAACTAGGCATTAATAAATAAATAATGATGACATTCACAGCAATCTAGATGAAATGG	2100
Db	55074	ACAACTAGGCATTAATAAATAAATAATGATGACATTCACAGCAATCTAGATGAAATGG	55133
OY	2101	AGACCTTATTTCTAAGTGGGGTAACTCAGAGATGAAAAACAAATCATATGTTCTCAC	2160
Db	55134	AGACCTTATTTCTAAGTGGGGTAACTCAGAGATGAAAAACAAATCATATGTTCTCAC	55193
OY	2161	TTACAAGTGGGGGCTAAGCTGTAGAGACACGAAGGATGGAATGATATATGAATCTGG	2220
Db	55194	TTACAAGTGGGGGCTAAGCTGTAGAGACACGAAGGATGGAATGATATATGAATCTGG	55253
OY	2221	GGACTTGAAGGGAAGATGGAAGAAGGCGAGGATTAAGACTACACAAATGGGTACAGT	2280
Db	55254	GGACTTGAAGGGAAGATGGAAGAAGGCGAGGATTAAGACTACACAAATGGGTACAGT	55313
OY	2281	GTACACTGCTCAGGTGATGGGTGACCAAAATCTCAAAATTCACATGAAGACTTATC	2340
Db	55314	GTACACTGCTCAGGTGATGGGTGACCAAAATCTCAAAATTCACATGAAGACTTATC	55373
OY	2341	CATGAAGCAACAACCACTGTTCCCAAAATCCCAATGAATTAATAATATATATA	2400
Db	55374	CATGAAGCAACAACCACTGTTCCCAAAATCCCAATGAATTAATAATATATATA	55433
OY	2401	ATGATTTTAAATTCACAGAAATTTAAAAAAGTTCACGTGTGAGATTATATATGAAGTA	2460
Db	55434	ATGATTTTAAATTCACAGAAATTTAAAAAAGTTCACGTGTGAGATTATATATGAAGTA	55493
OY	2461	AGAAATGAAGAAAGTGTAGCAAGTGTGAGCTCTGGAACAATGGACTTGAATTTTCACTTG	2520
Db	55494	AGAAATGAAGAAAGTGTAGCAAGTGTGAGCTCTGGAACAATGGACTTGAATTTTCACTTG	55553
OY	2521	CATACACTCTCTGCGCAATTTGGAAAAAGTATACACATGAATATATCAACCTATGATA	2580
Db	55554	CATACACTCTCTGCGCAATTTGGAAAAAGTATACACATGAATATATCAACCTATGATA	55613
OY	2581	AAGAAAAATCAAAAAAATTTGTGCAGGCACTTGTGACCTTGAAATGGTCCCATGATCTAC	2640
Db	55614	AAGAAAAATCAAAAAAATTTGTGTGAGCCACTTGTGACCTTGAAATGGTCCCATGATCTAC	55673
OY	2641	TTTTTCAATTTGGATATTAAGCCTCATATGATGATTTCACTTGTCTTAATGTGATGCTAG	2700
Db	55674	TTTTTCAATTTGGATATTAAGCCTCATATGATGATTTCACTTGTCTTAATGTGATGCTAG	55733
OY	2701	GCCCATATTTGATTTTAAATCAGGACAGCAATTAATTCTTAAGAGATTTGAACAAGATGG	2760
Db	55734	GCCCATATTTGATTTTAAATCAGGACAGCAATTAATTCTTAAGAGATTTGAACAAGATGG	55793
OY	2761	GACGTGATAGGAGAGGCTTAAATGTATGTGATATGGAACAAGGCCAAGATATCATTTGAG	2820
Db	55794	GACGTGATAGGAGAGGCTTAAATGTATGTGATATGGAACAAGGCCAAGATATCATTTGAG	55853
OY	2821	TTAGGATTTGTGTCTCAAAATACCTCTGGCCTCTGATTTTGGCCATAGTCTCTACAGAA	2880
Db	55854	TTAGGATTTGTGTCTCAAAATACCTCTGGCCTCTGATTTTGGCCATAGTCTCTACAGAA	55913
OY	2881	ATTAACAAGACTGTCTCAGCATCTTGTGTGATGCTGATTTGCTCACAGCTTTCAATTCAGT	2940
Db	55914	ATTAACAAGACTGTCTCAGCATCTTGTGTGATGCTGATTTGCTCACAGCTTTCAATTCAGT	55973
OY	2941	CCTGTAAGGCATCTCTGAATTTAAGCAACAAGAAAAGTCCCTGGAAGTCACTGAATCCCA	3000
Db	55974	CCTGTAAGGCATCTCTGAATTTAAGCAACAAGAAAAGTCCCTGGAAGTCACTGAATCCCA	56033
OY	3001	GAAGAGCTCTTACTTTAGCAACAAGGAGGTCTTCAACCTGGAACAAAGAAAGCAAT	3060
Db	56034	TAAAGGCTCTTACTTTAGCAACAAGGAGGTCTTTTACCACTGGAACAAAGAAAG-ACGAT	56092
OY	3061	AAGGTAAGTACAAAGACTCTCTTCTTCCACAGTCAAGTATGATTTTTGCTGTAAATTC	3120

Db 56093 AAGGTAAGTA-CAGAACTCTCTTCCACAGTCAATTATTTTGTGTAAGATC 56151
 Qy 3121 ATGTCCTTAATGCTTCCACCTTGTGTCTACATGAGGGGGTCAAGAGCTTTTCAGAGAA 3180
 Db 56152 ATGTCCTTAATGCTTCCACCTTGTGTCTACATGAGGGGGTCAAGAGCTTTGTGTAAGAA 56211
 Qy 3181 AGACAGAGACATGAAGCTTCTTTCAGAACTGAGTGTCTCAACCCAACTGTGTAG 3240
 Db 56212 AGACAGAGACATGAAGCTTCTTTCAGAACTGAGTGTCTCAACCCAACTGTGTAG 56271
 Qy 3241 CTCCTAAATGAGTCCCTCTTAATTTATCTCCCATCACTCTCTCAATCAATCA 3300
 Db 56272 CTCCTAAATGAGTCCCTCTTAATTTATCTCCCATCACTCTCTCAATCAATCA 56331
 Qy 3301 TTCATCTGCTCATGAGAGACCTGCTCTTACATTTAAAGAGAGAGAG 3360
 Db 56332 TTCATCTGCTCATGAGAGACCTGCTCTTACATTTAAAGAGAGAGAG 56391
 Qy 3361 ATGAGAGCATTTCTTCCAGAGTGAACCTGTATAGCCAGTTCTTTGTCACTTAC 3420
 Db 56392 ATGAGAGCATTTCTTCCAGAGTGAACCTGTATAGCCAGTTCTTTGTCACTTAC 56451
 Qy 3421 TTTTCAATTTATCTCTTGTCTATCTGGAAGGTCTAAGAGATATATAGTGGCCA 3480
 Db 56452 TTTTCAATTTATCTCTTGTCTATCTGGAAGGTCTAAGAGATATATAGTGGCCA 56511
 Qy 3481 ATATTAAGAGTGTTCATGAGAGAAATTTTAAAGATGACAGAGTTAAGGCTCAG 3540
 Db 56512 ATATTAAGAGTGTTCATGAGAGAAATTTTAAAGATGACAGAGTTAAGGCTCAG 56571
 Qy 3541 GATTCCTAAGAGCATATATGAGAGAGCATCTTCTCTCTTGTGCTGAACGAGACAG 3600
 Db 56572 GATTCCTAAGAGCATATATGAGAGAGCATCTTCTCTCTTGTGCTGAACGAGACAG 56631
 Qy 3601 GAAGAGACAGTATACATTTGTGCTGACATAGCCATGAGCATATAGGAGGAGAGCAT 3660
 Db 56632 GAAGAGACAGTATACATTTGTGCTGACATAGCCATGAGCATATAGGAGGAGAGCAT 56691
 Qy 3661 GAGAGAGCATATGAGAGAGCTGCTGCTGACAGAGCATATAGGAGGAGAGCAT 3720
 Db 56692 GAGAGAGCATATGAGAGAGCTGCTGCTGACAGAGCATATAGGAGGAGAGCAT 56751
 Qy 3721 TGGATCTGTGTTCCACCAAAATCTCATGTTGTAATTTTCCAAATGTTGAGAGAGG 3780
 Db 56752 TGGATCTGTGTTCCACCAAAATCTCATGTTGTAATTTTCCAAATGTTGAGAGAGG 56811
 Qy 3781 CCTTGTGAGAGATATATAGATCAGGGGATGTTTGGATGAATGTTTAAACACAT 3840
 Db 56812 CCTTGTGAGAGATATATAGATCAGGGGATGTTTGGATGAATGTTTAAACACAT 56871
 Qy 3841 CCCCCTTGTGATTTGTTGTGTAAGTCAAGAGTTCATGAATCTAGTTGTTTAAA 3900
 Db 56872 CCCCCTTGTGATTTGTTGTGTAAGTCAAGAGTTCATGAATCTAGTTGTTTAAA 56931
 Qy 3901 GCGGTGAGACCTCTCCCTCTCTTAATCTCTGCTCTACCATGTGAGAGCGCTCGCTC 3960
 Db 56932 GCGGTGAGACCTCTCCCTCTCTTAATCTCTGCTCTACCATGTGAGAGCGCTCGCTC 56991
 Qy 3961 CCCCCTTGTGATTTGTTGTGTAAGTCAAGAGTTCATGAATCTAGTTGTTTAAA 4020
 Db 56992 CCCCCTTGTGATTTGTTGTGTAAGTCAAGAGTTCATGAATCTAGTTGTTTAAA 57051
 Qy 4021 TATGCTTCTGTACAGTCTGTAGAGCTATTAAGCAGTTAAACCATTTCTTCAATAAT 4080
 Db 57052 TATGCTTCTGTACAGTCTGTAGAGCTATTAAGCAGTTAAACCATTTCTTCAATAAT 57111
 Qy 4081 TCCAGAGTCAAGAGTATTTCTTTTAAAGATTTGAATGAATCTAATACAGAGAGAG 4140
 Db 57112 TCCAGAGTCAAGAGTATTTCTTTTAAAGATTTGAATGAATCTAATACAGAGAGAG 57171
 Qy 4141 CCAGAGATGAATTCAGAGTCTTCTGCTGTCTTCAAGTCTCTGCTGTGCTC 4200
 Db 57172 CCAGAGATGAATTCAGAGTCTTCTGCTGTCTTCAAGTCTCTGCTGTGCTC 57231

Qy 4201 CCAGTGTCTCAATTCACCCAGAAACCAAGAAATTAAGAAATCCCATCTGATGTGTACATA 4260
 Db 57232 CCAGTGTCTCAATTCACCCAGAAACCAAGAAATTAAGAAATCCCATCTGATGTGTACATA 57291
 Qy 4261 GAAGCCACTCTTGGAGATGTCAAAACGAGTAAAGAAATGGAAGCAATCTCATGG 4320
 Db 57292 GAAGCCACTCTTGGAGATGTCAAAACGAGTAAAGAAATGGAAGCAATCTCATGG 57351
 Qy 4321 TAAATGAGATATCCCTCTGACCTTCTTGTATCTCTCTTAATTCCTGAGGCTTCTCATC 4380
 Db 57352 TAAATGAGATATCCCTCTGACCTTCTTGTATCTCTCTTAATTCCTGAGGCTTCTCATC 57411
 Qy 4381 TGAATGATCCCTGTCTCATTTTCAAGCTCTATCAAGTCTAATGTTTGGCTGTCTTTC 4440
 Db 57412 TGAATGATCCCTGTCTCATTTTCAAGCTCTATCAAGTCTAATGTTTGGCTGTCTTTC 57471
 Qy 4441 TCTACTGTCACTTTTAAAGAAATGTTTGCATTTGTTAAAGATGATAGAAATTAAT 4500
 Db 57472 TCTACTGTCACTTTTAAAGAAATGTTTGCATTTGTTAAAGATGATAGAAATTAAT 57531
 Qy 4501 GTAAATTTAAAGAAACATATGATTTTGTGTAAGATATAGTTGCTGATCTAATPAA 4560
 Db 57532 GTAAATTTAAAGAAACATATGATTTTGTGTAAGATATAGTTGCTGATCTAATPAA 57591
 Qy 4561 GACATGAAGAGA 4573
 Db 57592 GACATGAAGAGA 57604

RESULT 8
 AC139482/c 165490 bp DNA linear HTG 04-FEB-2003
 LOCUS Homo sapiens chromosome 5 clone Rpl1-583P24, WORKING DRAFT
 DEFINITION
 SEQUENCE: 12 unordered pieces.
 AC139482
 AC139482.1 GI:28201512
 HTG; HTG_PHASE1; HTG_DRAFT; HTG_ACTIVEPIN.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (04-FEB-2003) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 COMMENT
 -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: http://www.jgi.doe.gov

 Project Information
 Center Project Name: 1518520
 Center clone name: RPl1-1_583P24

Summary Statistics
 Consensus quality: 147035 bases at least Q40
 Consensus quality: 155210 bases at least Q30
 Consensus quality: 158567 bases at least Q20
 Estimated insert size: 175000; agarose-fp estimation
 Estimated insert size: 164390; sum-of-contigs estimation
 Quality coverage: 3.94 in Q20 bases; agarose-fp estimation
 Quality coverage: 4.19 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1249: contig of 1249 bp in length
 * 1250 1349: gap of unknown length
 * 1350 3680: contig of 2331 bp in length
 * 3780: gap of unknown length
 * 3781 6668: contig of 2888 bp in length
 * 6669 6768: gap of unknown length
 * 6769 10866: contig of 4098 bp in length
 * 10867 10966: gap of unknown length
 * 10967 17019: contig of 6053 bp in length
 * 17020 17119: gap of unknown length
 * 17120 22433: contig of 5314 bp in length
 * 22434 22533: gap of unknown length
 * 22534 34645: contig of 12112 bp in length
 * 34646 34745: gap of unknown length
 * 34746 49173: contig of 14428 bp in length
 * 49174 49273: gap of unknown length
 * 49274 67058: contig of 17785 bp in length
 * 67059 67158: gap of unknown length
 * 67159 84928: contig of 17770 bp in length
 * 84929 85028: gap of unknown length
 * 85029 114564: contig of 29436 bp in length
 * 11465 114565: gap of unknown length
 * 114565 165490: contig of 50926 bp in length.

FEATURES
 source
 1.165490
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="RP11-583F24"
 /clone_lib="RPCT human BAC library 11"

ORIGIN

Query Match 51.3%; Score 4424.6; DB 2; Length 165490;
 Best Local Similarity 87.3%; Pred. No. 0;
 Matches 5218; Conservative 0; Mismatches 634; Indels 123; Gaps 29;

2713 TTTTAAATCAGACAGCAATTACTTACAGAAAGTTGAACAAGATGGACGTGATAGGA 2772
 Db TCTCTCAAACTGCGCACCTGTTTATCAAGCAAGTTGAACAAGATGGATGATGGGA 79598
 2773 GAGGCTTAATGATCGGATATGCGACAGAGGCCAAGAAATCACTCAAGTTAGATTTGG 2832
 Db GAGGCTTAATGATCGGATATGCGACAGAGGCCAAGAAATCACTCAAGTTAGATTTGG 79538
 2833 TCTCAATACCTCTGGCTCTGATTTGCCATAGTCTCATAGAGAAATTAACAAGCTG 2892
 Db TCTCAATACCTCTGGCTCTGATTTGCCATAGTCTCATAGAGAAATTAACAAGCTG 79478
 2893 TCCAGATTTGTAAGCCTGATGCTCTCAACAGCTTTCATTTAGCTCTGTAGGCACT 2952
 Db TCCAGATTTGTAAGCCTGATGCTCTCAACAGCTTTCATTTAGCTCTGTAGGCACT 79418
 2953 TCCGAATTAAGCAACAGAAAGTCTCTGGAAGTCACTGAATCCAGAAAGGCTCTCT 3012
 Db TCCGAATTAAGCAACAGAAAGTCTCTGGAAGTCACTGAATCCAGAAAGGCTCTCT 79358
 3013 ACCTTAGACAGAGGAGGCTTCCACCACTGACCAAGAAAG--GAACGATPAGGTAAGTA 3071
 Db ACCTTAGACAGAGGAGGCTTCCACCACTGACCAAGAAAGTAAGGTAAGTA 79298
 3072 CCAAGAACTCTCTTCTTCAACAGTCAAGTATGATTTTGTGTAGATCAATGCTTTATG 3131
 Db CCAAGAACTCTCTTCTTCAACAGTCAAGTATGATTTTGTGTAGATCAATGCTTTATG 79242
 3132 CTTTCAACCTTGTGCTACATGACGGGGGTCACAGAGCTGTTTTCAGAAAGAAACAGAGAC 3191
 Db CTTTCAACCTTGTGCTACATGACGGGGGTCACAGAGCTGTTTTCAGAAAGAAACAGAGAC 79241
 3192 ATGAAGCTTCTTTTCAAGAACTGAGTGTCTCAACCAACTGTGTAGCTTAAATGCT 3251

79181 ATGAAGTTCTCTTTCGAAACAGAGTGTGCGCAACCCAGACTGTGTGATTCACACTGGA 79122
 3252 GTCCCCCTTTCAATTTATCTCCCATATCACTCTCTTCAATTCATTCATCTGCC 3311
 Db GTCCCCCTTTCAATTTATCTCTCCCATATCACTCTCTTCAATTCATTCATCTGCC 79062
 3312 CTCATGAGAGACTGTGCTCTTCAATTCATTCATTCATTCATTCATTCATTCATTCATTC 3371
 Db CTCATGAGAGAGTGTGCTCTTCAATTCATTCATTCATTCATTCATTCATTCATTCATTC 79002
 3372 TCTTCCAGAGTTGAACCTGCTAG--AGCAGTTTCTTGTTCATCTTCAATTCATTCATTC 3430
 Db GCTTCAACAGTTGAACCTGCTAGAGAGGTTTGTGTTCATTCATTCATTCATTCATTCAT 78942
 3431 TATTTCTTTTGTCTATCTGAAAGGCTTAAGAAAGATATAGATGCCCC-----ATA 3483
 Db TATTTCTTTTGTCTATCTGAAAGGCTTAAGAAAGATATAGATGCCCC-----ATA 78882
 3484 ATTAAGAGTGTTCATGAGAAAGTATTAACAAAGATGACAGAGTTAA----- 3533
 Db ACTGAGAGTGTTCATGAGAAAGTATTAACAAAGATGACAGAGTTAA----- 78882
 3534 -----GGTTCAGAGTCTTAAGACAAATATATAGGGAGCACTACTTCTCCCTAG 3585
 Db -----GGTTCAGAGTCTTAAGACAAATATATAGGGAGCACTACTTCTCCCTAG 78762
 3586 GCTGAAACGACAGAGGAGAGACAGTTCATTTGCTGCTATGCTTATGCTTATGCTTATG 3645
 Db GCTGAAACGAGAGAGAGAGACAGTTCATTTGCTGCTATGCTTATGCTTATGCTTATG 78702
 3646 AGGAGGAGAGATAGAGGAGGCAAGTGAAGGCTGCGGCAAGCAGCAGCAGCA 3705
 Db AGGAGGAGAGATAGAGGAGGCAAGTGAAGGCTGCGGCAAGCAGCAGCAGCA 78642
 3706 CAGGCTGATATAGTTGATCTGTGTCCACCAAAATCTCATGTGATGTATTTTCA 3765
 Db CAGGCTGATATAGTTGATCTGTGTCCACCAAAATCTCATGTGATGTATTTTCA 78582
 3766 ATGTTGAGAGAGGCTGTTGAGAGATATTTAGATCAGGGAGTGTTCAGATGA 3825
 Db ATGTTGAGAGAGGCTGTTGAGAGATATTTAGATCAGGGAGTGTTCAGATGA 78581
 3826 ATGTTTAAACATATCCCTTGTGATTTGTTGATGATGACAGATTCATGAA 3885
 Db ATGTTTAAACATATCCCTTGTGATTTGTTGATGATGACAGATTCATGAA 78524
 3886 TCTAGTTGTTAAAGCGTGAACCTCTCC--TCTCTTACCTCTGCTACCA 3943
 Db TCTAGTTGTTAAAGCGTGAACCTCTCC--TCTCTTACCTCTGCTACCA 78467
 3944 TGTGAGAGCGTCTGCTCCCTTGTGCTTACACAGATGGAACCTCTGAGGCTCC 4003
 Db TGTGAGAGCGTCTGCTCCCTTGTGCTTACACAGATGGAACCTCTGAGGCTCC 78407
 4004 CCAAGAGAGAGCTGATGCTCTTGTGACAGTGTGAGAGCTTATGAGCAAGTAAAC 4063
 Db CCAAGAGAGAGCTGATGCTCTTGTGACAGTGTGAGAGCTTATGAGCAAGTAAAC 78347
 4064 CATTTCTTCAATAATTTCCAGTCAAGTATTTCTTTTGAATTTGAAATGAACT 4123
 Db CATTTCTTCAATAATTTCCAGTCAAGTATTTCTTTTGAATTTGAAATGAACT 78287
 4124 AATACAGACAGAGGAGGAGATGAAATCCCAAGTCTTCTGCTGCTCTCAG 4183
 Db AATACAGAGGAGGAGGAGGAGATGAAATCCCAAGTCTTCTGCTGCTCTCAG 78227
 4184 TCTCCGCTGTGTCTCCAGTGTCTCAATTCACACAGAAACAGAAATTAAGAAATCC 4243
 Db TCTCCGCTGTGTCTCCAGTGTCTCAATTCACACAGAAACAGAAATTAAGAAATCC 78167
 4244 CACTGATGTGTATAGAGCACTCTTGGAGTGTCAACAGATTAAGAAATG 4303

Db 78107 CACTGATGTGTGATAGAAAGCCACTCTTGGGGTGTCAAAACAGATAGAAAGATGG 78048
 Qy 4304 AAGCAATCTCTCATGTGTAAATGAGACTATCCCTCTCACTTCTTGTATCTCTTAATTC 4363
 Db 78047 AAGCAATCATCATAGCAAAATGAGACTATCCCTCTCTCTTGTATATCTCTTAATTC 77988
 Qy 4364 CTGGGGCTTCTCATC-----TGATTGATCCCTGTCTCATTTCAAGTCTATAGACTA 4417
 Db 77987 CTGGGACAGTCTGTCTACAGGTGATGATTCCTGTCTCATTTCAAGTCTTTCAGACTA 77928
 Qy 4418 CTTTAATGTTGGCTGTCTTCTCTACTGTCTACTTATGAGAAATGTTGCAATTGT 4477
 Db 77927 CTTTAATGTTGGCAATGCTTCTCTACTGTCTACTGCAAAATGTTTGCATTCAT 77868
 Qy 4478 TAAAAATGATAGAAATTAATATTTTAAAA--AGAACATATGATTTTGTTTAGA 4535
 Db 77867 CCAAAATGATAGAAATTAATATTTTAAAAAGAGAACATATTTCTTTGTTAGA 77808
 Qy 4536 ATATATGTTGGCTGATCTTAATAAGACATGAGAAAGAAATTTCTTAACAGAAAGTAT 4595
 Db 77807 ATACAAATTTGACCCCTCTCTAATAAGACAAAGTATGAAATTTTAAATTAAGAAATAC 77748
 Qy 4596 AGTTGCTCTCTGGGTCACTAGTTCTGAATCTACAGATTCAAGAACTACAGAGAGAA 4655
 Db 77747 AGTGTGCTCTCTGGGTCACTAGTCTGTGAATCTGTGAATCTCAACCCACAGGTGAAA 77688
 Qy 4656 CTTTTCAGAAATTAAGGTGTGGGAGTTGTGTATGACTGAACAGGTACAACTTGTGA 4715
 Db 77687 CTGTTCAGAAATTAAGGTGTGGGAGTTGTGTATGACTGAACAGGTACAACTTGTGT 77628
 Qy 4716 TTTCTTGTCTATTTTCTGAAAACTACATATTAACAAAGCTTATATAGCATTTGCTAT 4775
 Db 77627 TTTCTTGTCTATTTTCTGAAAACTACATATTAACAAAGCTTATATAGCATTTGCTAT 77568
 Qy 4776 TTTGTCACTATTTCTTAATAACT--TAAATGATTTAATGTCTGGGAGAAAGTCAATAG 4834
 Db 77567 TTTGTCACTATTTCTTAATAACTTTAGAAAGTATTTAATGTCTGGGAGAAAGTCAATAG 77508
 Qy 4835 AGTATATACAAATACCA-----TATATAGAAATTTAGCATCTGACATTTTGGTCT 4886
 Db 77507 GTTTTGTACAACTATCATCATTTTATATTAAGAAATTTAGCATCTGACATTTTGGTCT 77448
 Qy 4887 TGTGTGGGGTCTGAGAA-----GAATCCCTGTAAATTAACAAAA 4928
 Db 77447 TGTGTGGGGTGAAGACACACATCTGCCATATGCAAAAGCCCTGTGAATATCCAGAA 77388
 Qy 4929 ATGACATCTTTCGATCTGTAACTAGAAAGCTTCAAGCATCATCATAGAAATTCAGAA 4988
 Db 77387 ATGACATCTTTCGATCTGTAACTAGAAAGCTTCAAGCATCTCATAGAAATTCAGAA 77328
 Qy 4989 ATTGTGCTCCCAAGTTCTTAGAGAGTTGCCCTCATCTTGTGTACTTACATGCTTCCA 5048
 Db 77327 ACTGTACTCTCGAGTCCCTTAGAGAGTTGCCCTCATCTGTGTACTTATGCTTCCA 77268
 Qy 5049 GCGACATTAACATTCACAGTCTTAATGAAAAAGACAGAGGAGAAAGAGCTTGTGCTCT 5108
 Db 77267 GCGACATTCATTAACATTCAGATGAGAAAA--AAATAGAGAAAGAGATGCTGCTCT 77209
 Qy 5109 TCTATTAATCCCATGAGCAGAGACTTGTCTGTCTGTCACTTTGTGATTTCTTCACTTAACA 5168
 Db 77208 GCTATGTGTCTGTAGAGCAGAGACTTGTCTGTCTGTCACTTTGTGATTTCTTCACTTAACA 77149
 Qy 5169 GCACTGTCTCATGGAGTGTATCCAGACATCAAGGAAAACTGGGATGTGGTCTTGTGCT 5228
 Db 77148 GCACTGTCTCATTAATGAGCACTCAGACATCAAGGAAAACTGGATGTAGGTCTTGTGCT 77089
 Qy 5229 GCTTGTACATTTCTAGAAAGTTATGTGACCA-----AAAAGAAATCTGGGCAACC 5283
 Db 77088 GCTTGTAACTCTAGAAATTTTATGTGACCAACAGGAAAAATAAACACTGGGCAACCT 77029
 Qy 5284 AGCAGTCTCTTCAAGCCCTGACTGTCTGTATTTCTGTGCTCAGATCAAGATTTTTCAGGA 5343
 Db 77028 AGCAGTCTCTTCTGCC-----TGTGTCTGTATTTCTGTACTCATCAAGATTTTTCAGGA 76973

Qy 5344 ACTCTCAGAAATTAATAATGTTGGGCAAGAAACAGAACTGAGTCTGTGACAGACTC 5403
 Db 76972 ACTCTCTGAAATTAATAATATATGCGCAGAGAAACATATAGAAAGTCTCATGAGGGCTC 76913
 Qy 5404 CAGGACACAGGGGCTGTATTTGACCTGTCTTCAATGTTGAAAC--CAGGAAACCCCTT 5462
 Db 76912 CAGGACACAGGGGCTGTGACCTGTCTTGTGTGTGAACTCATGAAAACTTT 76853
 Qy 5463 AATTCTAGGCTTACCTTCACTTATGTTATATGAGATTAATACATAGACAGTCTTT 5522
 Db 76852 AATTCTAGGCTTACCTTCACTTATGTTATATTAAGATTAATACATAGATATCTTT 76793
 Qy 5523 AAGACATCATATGATGTTAAACACATGCTAAATGTTGGTATACACAGTAAAAAG 5582
 Db 76792 AAGAAC-----ACATCTAAATGTTGGTATACACAGTGAAG 76751
 Qy 5583 AAGGACATGACTTACCTTACGATCTTGGGTTATATGAGAGAAACAAATATACA-T 5641
 Db 76750 AAGGACATGAC-TAGCGTTATATGAGCTTCAAGTTTATATGAGAGAAACAAACACACTT 76692
 Qy 5642 ACATACCTATATGATGACAAACAGTTAGTCTCTGAGTGTGATTAACAGAGTCTCC 5701
 Db 76691 ACCCTAATATATAGAGACATATGTTTATGTTGCTGCTGCTGTGATTAACAGAGTCTCC 76632
 Qy 5702 TTTTCTCCCATTTCTTTTGGGCAATCAGACCTGTGCACTTCTCTTAAGAGA 5761
 Db 76631 TTTTCTCCCATTTCTTTTGGGCAATCAGACCTGTGCACTTCTCTTAAGAGA 76572
 Qy 5762 GCTCATATGATGATCTCATCTCTGTATGCTCTCTTATCTCCAGAGAGATGATCT 5821
 Db 76571 GCTCATATGATGATCTCATCTCTGTATGCTCTCTTATCTCCAGAGAGGATGATCT 76512
 Qy 5822 TCTTTCACCTGAGAGCTCTGCGCAATGATCTTCTTGGATTTCCAGAGAAAGTGGC 5881
 Db 76511 TCTTTCACCTGAGAGCTCTGCGCAATGATCTTCTTGGATTTCCAGAGAAAGTGGC 76452
 Qy 5882 CTGTATAGGCAAAAAAGAACTCTGTAATTTTCTTAATATGACGCACTCACTCTAT 5941
 Db 76451 CTGTATAGGCAAAAAAGAACTCTGTAATTTTCTTAATATGAGCTTCACTCTCTAT 76392
 Qy 5942 TTT--TCCCTATTTTCACTTCTCTCATCTCATCTGAGATTTGTTAGTTA--ATT 5997
 Db 76391 TTTGTTCTCTATTTGACCTCTCTCTCTCTCATCTGAGATTTGTTAGTTA--ATT 76332
 Qy 5998 TTTTTCAGCCCAATTTTGAAGTCACTGATTTTAACTTGAATCACTCTCT 6057
 Db 76331 TTTTTCAGCCCAATTTTGAAGTCACTGATTTTAACTTGAATCACTCTCT 76272
 Qy 6058 ACTTACCCCTCTTAACATGATTAATGACACATATGAGTGTGGTCCAAAGGCTCG 6117
 Db 76271 CTTTACCCG--CTCTAATCATGATTAATGACACATATGAGTGTGGTCCAAAGGCTCG 76213
 Qy 6118 TGAATAAATGAGATGATGAGTCAAGCCCTGTGGGCTCACTTCACTATTAATTAATTA 6177
 Db 76212 TGAATAAATGAGATGATGAGTCAAGCCCTGTGGGCTCACTTCACTATTAATTAATTA 76160
 Qy 6178 CCCCCGGAATAATATATCTCTTATATCTTATATTTCACTATTCGCAATTCAGAAA 6237
 Db 76159 GTCGCAATAATATATCTCTTGTGTGTCAA--TGCAATATCTGACATATGAGAA 76101
 Qy 6238 TAAATCTAATTTTACCTCTAGGCTCTTCAATGATTAATAAGGCAATATCTAATTA 6297
 Db 76100 TAAATCTAATTTTACCTCTAGGCTCTTCAATGATTAATAAGGCAATATCTAATTA 76041
 Qy 6298 CTGTGATCAAGTCTGTTATTTCTCCCAATTTGACCTATTCCTCACTTCCAGTTGAAA 6357
 Db 76040 CTGTGATCAAGTCTGTTATTTCTCCCAATTTGACCTATTCCTCACTTCCAGTTGAAA 75981
 Qy 6358 TTTGATGAATATCTCTTTTAAACCAAGCCCTACATCTTCTTCACTTCACTGAG 6417
 Db 75980 TCTACATGAATATCTCTTTTATCTCAAGCCCTACATCTTCTTCACTTCACTGAG 75921

Oy 6418 ACTCTCTCATCAAAATGTTGGATCAACAAAGAAACGCTACCAAGATCTCCGAAAG 6477
 Db 75920 ACTCTCTCATCAAAATGTTGGATCAACAAAGAAACG-AAACAAAGATCTCTGAAAG 75862
 Oy 6478 GAGAAAGAAATGAGTTTACATTTGTATTAAGTCAAGAAACACTTAAAGTCCCATAGA 6537
 Db 75861 AAGAAAGAAATGAGTTTATTTGTATTAAGTCAAGAAATCACTTAAAGTCCCATAGA 75802
 Oy 6538 TATTTCCCACTCTCAATTAAGTCTGATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 6597
 Db 75801 TATTTCCCACTCTCAATTAAGTCTGATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 75742
 Oy 6598 CCTCTCTCTATTTGATGAGCAATTCAGTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 6656
 Db 75741 CCTCTCTCTATTTGATGAGCAATTCAGTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 75682
 Oy 6657 CAAAATGCTTCTAGGGGTTCCGCAAAAGCAAGCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 6716
 Db 75681 CAAAATGCTTCTAGGGGTTCCGCAAAAGCAAGCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 75622
 Oy 6717 TGTCAATTTGAGTTTGGGCACTCAATCAATCTTCTGCAAGCAAGTCAAGTCAAGTCAAG 6776
 Db 75621 TGTCAATTTGAGTTTGGGCACTCAATCAATCTTCTGCAAGCAAGTCAAGTCAAGTCAAG 75562
 Oy 6777 CACAATAGAAATTCATGATGAATGAATTTTCTCATCTAATTAATGATGATGATGATGAT 6836
 Db 75561 CACAATAGAAATTCATGATGAATGAATTTTCTCATCTAATTAATGATGATGATGATGAT 75502
 Oy 6837 GCTTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6896
 Db 75501 GCTTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 75442
 Oy 6897 GCTTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6956
 Db 75441 GCTTCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 75382
 Oy 6957 TATGAAATGAGCACTAGAGATTAATGAATGATGATGATGATGATGATGATGATGATGAT 7016
 Db 75381 TATGAAATGAGCACTAGAGATTAATGAATGATGATGATGATGATGATGATGATGATGAT 75322
 Oy 7017 CACTGATTTTCTCAAGCAAGTCCCGCATTTTCAAG-ATCTGAGGAAGATC 7074
 Db 75321 TACTGATTTTCTCAAGCAAGTCCCGCATTTTCAAG-ATCTGAGGAAGATC 75262
 Oy 7075 CCAATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7134
 Db 75261 AATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 75202
 Oy 7135 GTCATGAGGATGATCAAGCAAGTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 7194
 Db 75201 GTCATGAGGATGATCAAGCAAGTCTTGTGATGATGATGATGATGATGATGATGATGATGAT 75142
 Oy 7195 TTCAGGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7254
 Db 75141 TTCAGGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 75082
 Oy 7255 TTTGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7314
 Db 75081 TTTGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 75022
 Oy 7315 GCTTCAAGTCAAGCAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 7374
 Db 75021 GCTTCAAGTCAAGCAAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 74962
 Oy 7375 AGGTGATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7434
 Db 74961 AGGTGATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 74903
 Oy 7435 ACATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7494
 Db 74902 ACATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 74843
 Oy 7495 GTTTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7554

Db 74842 GTTTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 74783
 Oy 7555 ACAGAACTGACACCAATCAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 7614
 Db 74782 ACAGAACTGACACCAATCAAGGATGATGATGATGATGATGATGATGATGATGATGAT 74723
 Oy 7615 TTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7674
 Db 74722 TTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 74663
 Oy 7675 TGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7734
 Db 74662 TGTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 74603
 Oy 7735 GCGGCGATTTCTCTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 7794
 Db 74602 GCGGCGATTTCTCTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 74543
 Oy 7795 ATCCGCAATCCCATCTCCAAATTCCTCAATCTGATGATGATGATGATGATGATGATGATGAT 7854
 Db 74542 ATCCGCAATCCCATCTCCAAATTCCTCAATCTGATGATGATGATGATGATGATGATGATGAT 74483
 Oy 7855 CTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7914
 Db 74482 CTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 74423
 Oy 7915 TACCAATGCGGCG 7974
 Db 74422 TACCAATGCGGCG 74363
 Oy 7975 TCCCTGCTCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8034
 Db 74362 TCCCTGCTCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 74303
 Oy 8035 TCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8094
 Db 74302 TCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 74243
 Oy 8095 GTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8154
 Db 74242 GTTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 74183
 Oy 8155 CCGCTGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8214
 Db 74182 CCGCTGACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 74123
 Oy 8215 CTGCTCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8274
 Db 74122 CTGCTCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 74063
 Oy 8275 TTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8334
 Db 74062 TTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 74003
 Oy 8335 ATCAATTAATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8394
 Db 74002 ATCAATTAATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 73943
 Oy 8395 GTTCTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8454
 Db 73942 GTTCTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 73883
 Oy 8455 CAGGAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8514
 Db 73882 CAGGAAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 73823
 Oy 8515 TCAGACAGACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8574
 Db 73822 TCAGACAGACTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 73763
 Oy 8575 TTAGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8609

DB 73762 TTAGCCTTGGCCAGAAATGCTCAGTGGTTCC 73728

RESULT 9
AC103974
LOCUS
DEFINITION Homo sapiens chromosome 11, clone RP11-1081L13, complete sequence.
ACCESSION AC103974
VERSION AC103974.6
KEYWORDS GI:20531887
HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 189230)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 11, clone RP11-1081L13
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 189230)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Baetien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamasares, R., Landers, T., Lehoczyk, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meltrin, J., Menue, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 189230)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Baetien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamasares, R., Landers, T., Lehoczyk, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meltrin, J., Menue, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (26-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 4 (bases 1 to 189230)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Baetien, V., Bloom, T., Boguslavsky, L.,

REFERENCE 5 (bases 1 to 189230)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Baetien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamasares, R., Landers, T., Lehoczyk, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meltrin, J., Menue, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (13-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 13, 2002 this sequence version replaced gi:20330974.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L21856
Center clone name: 1081_L13

FEATURES
source

Location/Qualifiers
1. 189230
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-1081L13"
/clone_1lb="RPC1-11 Human Male BAC"
198. 223
/rpt_family="CTCA n"
1203. 1215
/note="<30 qual SNGL region"
2295. 2378
/rpt_family="L2"
complement(2381. 3394)
/rpt_family="L1MD2"
3396. 4062
/rpt_family="L2"
4172. 4245
/rpt_family="AT_rich"
4411. 4465
/rpt_family="TA n"
complement(4482. 4546)
/rpt_family="L1MC4a"
complement(4547. 4857)
/rpt_family="AluJo"
complement(4858. 5400)
/rpt_family="L1MC4a"
6131. 6239
/rpt_family="L2"
complement(6654. 7008)
/rpt_family="WL71B"
7551. 7749
/rpt_family="L2"
8025. 8057
/rpt_family="AT_rich"

	repeat_region	9218..9800	/rpt_family="L1PB1"
	repeat_region	9883..9955	/rpt_family="L1PB1"
	repeat_region	complement(11450..12969)	/rpt_family="L1PA4"
	repeat_region	complement(13902..14364)	/rpt_family="LTR33"
	repeat_region	complement(14660..15009)	/rpt_family="MER2"
	repeat_region	complement(15115..15369)	/rpt_family="MLT1J"
	repeat_region	15517..15603	/rpt_family="MIR3"
	repeat_region	complement(15612..15692)	/rpt_family="L2"
	repeat_region	16230..16337	/rpt_family="MIR"
	repeat_region	complement(16959..17034)	/rpt_family="L2"
	repeat_region	17340..17464	/rpt_family="MIR"
	repeat_region	19511..19704	/rpt_family="LMB5"
	repeat_region	19705..19813	/rpt_family="LMB4"
	repeat_region	20596..20683	/rpt_family="L2"
	repeat_region	complement(20686..20691)	/note="<30 qual SNGI_region"
	repeat_region	complement(20691..20874)	/rpt_family="LMD3"
	repeat_region	21261..21416	/rpt_family="MIR"
	repeat_region	complement(22006..22307)	/rpt_family="L1P4"
	repeat_region	complement(22308..22559)	/rpt_family="Alusg1"
	repeat_region	complement(22560..24537)	/rpt_family="L1P4"
	repeat_region	24537..24718	/rpt_family="L1"
	repeat_region	complement(24717..25049)	/rpt_family="L1P4"
	repeat_region	complement(25039..25908)	/rpt_family="L1PA4"
	repeat_region	complement(25985..27145)	/rpt_family="L1"
	repeat_region	complement(27179..29623)	/rpt_family="L1P4"
	repeat_region	complement(29632..30526)	/rpt_family="L1"
	repeat_region	complement(30530..31245)	/rpt_family="L1MEC"
	repeat_region	complement(31246..31555)	/rpt_family="Alusx"
	repeat_region	complement(31556..31800)	/rpt_family="L1MEC"
	repeat_region	complement(31853..32053)	
Query Match		30.2%; Score 2601.8; DB 9; Length 189230;	
Best Local Similarity		72.3%; Pred. No. 0;	
Matches 4441; Conservative		0; Mismatches 1072; Indels 626; Gaps 57;	
OY	2591 CAAAAATTGTGTCAGGCCATTTGCAGCCTTGGAATGATGATCTACTTTTCATT	2650	
Dd	13023 CACAGATGGGATCAACCCATTTGAAAGCCCGAAGAACCCTATGATCTTAATTAGTTTT	13082	
OY	2651 GGATATAAGCCCTCATATGATGATGTCACATTTGCTT-AATGATGCTTAGGCCCATAT	2709	
Dd	13083 AGTTATTAAGCCTCACAAAACAAGCTCAGCTTGCTTGACATGATTAACCTAGGCCATAT	13142	
OY	2710 TGAATTTTAA-ATCAGACAGCAATTACTTACAGGAAGTTGAMCAGATGGGACGTAT	2768	

Db	13143	AAATTGMAAAAGTTCCAAAACAGCAATTACTTACACGAAAGTTGAATGAGATGGGAATATGAT	13202
OY	2769	AGGAGAGGCTTAAATATGTACTGTGATATGGACAGAGGCCAAGATCATCTCAGTTAGATT	2828
Db	13203	GAGAGAGGACAGATATGTATTAATAATGGGACAGAGGCCAAAGAAACATCTCAGTTAGGCAAT	13262
OY	2829	TGCTCTCAAAATACCTCTGGCCCTGTATTTGGCCCAATGCTCCATACAGAAATAACAG	2888
Db	13633	CGTGGCTCAAAACAGCTCTGGCTCTGATTTGGCTGTGAATCTCATACAGAAATCATCATG	13322
OY	2889	ACTGTCCAGCATCTTCGTAGCCCTGTGATTGCTCACCAGCTTCATTTCAGCTCTGTAGG	2948
Db	13323	--CTGTCTGTAATCTTCAATAAAGCTGGGTGTGGCCAGTTCTCATTTCCAGCTCTGTAGG	13380
OY	2949	CATTCCTGTAATTAAGCAACACAGAAAAGTCTCTGTAAGTCACTGAATCCCAAAAAGCTT	3008
Db	13381	CATCTCCTAAATATGACACACAGAAATATGTCCTGTAAATCTCAGAAAGAGCT	13440
OY	3009	CTCTACCTTTAGACACAGGAGGCTTTCACC-----ACTGGCAAAAGAAAGCAATAG	3063
Db	13441	CTACCCCTTTAGACAGGAGGAGCTTTTATCTTTGACCTTGAAGAAAGAGAAAGGTAAG	13500
OY	3064	GGTATAGTACCAAGAACTCTCTTC--TTCCACAGTCAAGTATGATTTTGTCTGTAAATCAT	3122
Db	13501	GGTATGTATCCAAAGAACTCTCTTCTTTCCACAGAAATATGTCTTGTCTGTATTAAGTTCAC	13560
OY	3123	GTCTTATGCTTCACCTT--GGTGTACATGCAAGGGGGTCAAGAGCTTGTTCAGAAA	3180
Db	13561	AAACCTGTGTATTCACCTTGGGGTTCTGTACATACAGAAAGTTCATTAATTTCTTCAGAAA	13620
OY	3181	AGACAGAGACATGAAGCTTCCTTCAGAAACAGAGGCTGTCAACCCAACTGTGTGAG	3240
Db	13621	AGCAGAGACATGAAGCTTCCTTCAGAAACAGAGGCTGTCAACCCAACTGTGTGAG	13680
OY	3241	CTCTAAATGATGTC-----CCCCCTCTTAATTTATCTCCCATATCACCTCTCTC	3290
Db	13681	TTTCAAAATGGGGTCAGCAGAGCTTTCTCCCTGTATTCATCTCTCATATCACCCCTTT	13740
OY	3291	ATTTCATATTCATATTCGCTCTCATGAGAGACTGTGCTCTTACATTCATTTAAGCA	3350
Db	13741	ATTTCATATTCATATTCGCTCTCATGAGAGATATATGTCTTTTAAATTCATTTAAGCA	13800
OY	3351	GCAAGGGGACATGAGGCAATTTCTCCACAGATTGAATGCTATGAGCCAGTTCTTGG	3410
Db	13801	GCAAGGGGATATGAGGCAATTTCTCCACAGATTGAATGCTATGAGCCAGTTCTTGG	13855
OY	3411	TTTCACTTACTTTCAAAATTTATTCCTTGTGCTATCTTGAAAGGCTTAAGAGATAT	3470
Db	13860	TTTATCTTACTTTCAAAATTTATTCCTTGTGCTATCTTGAAAGGCTTAAGAGATAT	13919
OY	3471	AGATGAGCC-----AATAATTAAGAGTTCATGAGAAAGATTTACAAAGATGC	3523
Db	13920	AGATGAGTCCAAATATGAATTAATGAAGAGTTCATGAGGAGCATATGATCAAAATATG	13979
OY	3524	ACAGAGTTAAAGG-----TCAGATCTTAAGCAGCAATATACATAGGGAGCA	3569
Db	13980	ACAGAGTTAAAGGAAAGCAATGCTTAGGAGAGCACTTAATGCAATGTGTAGCA	14039
OY	3570	CTACTTCTCCCTAGGCTGAACAGACAGGAAAGACATTTACATTTG-----TGCC	3623
Db	14040	CTACTTCTCCCTAGGCTTAAGAGGACAGGAAAGACATTTACATTTAGAGCACAGC	14099
OY	3624	CATAGCCATAGCTGTAGCCATTAAGGGTGGAGAGCATGACAGGCAATGAGAAAGCCTT	3683
Db	14100	CATAGCCATAGCTGTAGCTGTAGGGTGGAGAGATGACAGGCAAGGTGGAGAAAGCCTT	14155
OY	3684	GCGAGGCCAAGCAGCAGCCACACAGGCTGATATATTTGGATCTGTGTCCACCAAAT	3743
Db	14160	GCAATCACCAACACAGCCACAAAG-----	14188
OY	3744	CTCATGTGTAATTTCCAAATGTTGAGAGAGGCCCTTGTGGAGATGATTAATTAAG	3803

Db 14186 ----- 14185
 Qy 3804 TCACGGGAGATGTTTGCATGAAATGTTTAAACACCATCCCTTGGTATGTTGTG 3863
 Db 14186 ----- 14185
 Qy 3864 ATACTGACGAGTTCTCATGAAATCTAGTTTAAAGCGTGTAGCACTCTCCCTTC 3923
 Db 14186 ----- 14185
 Qy 3924 TCTTACTCTGCTCTCACCATGTGAGACGCTCGCTCCCTTGGCTTTGACAGAGAT 3983
 Db 14186 ----- 14185
 Qy 3984 GGAAGCTTCTGAGGCTCCCGAAGAGAGCTGTATGCTTCTGTAGACGTCTGTAG 4043
 Db 14186 ----- 14185
 Qy 4044 AGCTATTAGCCAGTTAAACCATTTCTTCATTAATTTCCAGTCTCAGTATTTCTTT 4103
 Db 14186 ----- 14185
 Qy 4104 TAGCAATTTGAGATGAATATAACAGAGAGAGCGAGAGATGGAATCCCAAGT 4163
 Db 14186 ----- 14215
 Qy 4164 GCTTTCCTGCTGTCTTCAGATCTCTGCTGTGTCTCCAGTGTCTCAATTCACAGAA 4223
 Db 14216 ACTGCTCTGCTGCTTCAGTCTCTACAGTGTCTCAGTGTCTCAGTTCAACAGAA 14275
 Qy 4224 ACCGAAATTAAGAAATCCCATGTATGTGTATAGAAAGCACTCTCTGGAGATGCA 4283
 Db 14276 ACCA-AAAAGAAAGAAATATCACTGATGAGTACAGAAAGCCCTCTGGAGATGCA 14334
 Qy 4284 AACAGATTAAGAAAGTGAAGAAAGCAATCTCATGTGTAATGAGATATCCCTCTCACC 4343
 Db 14335 AACAGGATGAAGAAATTTGGAGAGCAATCTCTCAAGGAAATGAAACATTCCTCACC 14394
 Qy 4344 TTTCTGTA-TCCTCTAATTTCTGAGGCTTTCTCT-ATCTGATGATCCCTGT 4394
 Db 14395 TTTTATATCTTCTCTACCTACCTCGGGGTTTTCTCTAATCTCATGAAATCATGT 14454
 Qy 4395 CTCATTTCACTCATACGA-CTACTTAAATGTTTGTGCTGTCTCTCTACTCTACTT 4453
 Db 14455 CTCATTTAATCTCATACGAATTTTTTTTATGTTGAAATCTTCTCTATGTGCGAT 14514
 Qy 4454 TTATGCAAGAAATGTTGATTTGTTAAATGATGATAGAAATTAATTTTAAATA 4513
 Db 14515 TTATGAGAAATTACTGATTTATCAAAAATGCTAGAGATGAAATTTTAAATA 14574
 Qy 4514 G---AACATATGATTTTGTGTAATTAAGTTGTGCTGATCTAATTAAGACATGAG 4569
 Db 14575 GAGAAACATATCTCTTTTGTGTAATTAAGTTGTGCTGCTCTAATAAAGAACGAG 14634
 Qy 4570 AAGAAATATCTTAAACAGAAAGATATGTTGTGCTGTGCTACTAGTCTGAATTA 4629
 Db 14635 TAGAAATATCTTAAACAGAAAGTACAGTTGTCTCTCATGTCAAGTTGTCTGATCC 14694
 Qy 4630 CAGATTCAAACAACTACAGAGAGAACTTTTCCAAAATA-AAGTGTGGGAGGTTGTG 4688
 Db 14695 TNGATTCACAAACACAGAGTGAATATTTCCAAAAAAGAGACAGAGTACGTGA 14754
 Qy 4689 TATGTACTGAACAGATCAAACTGTATTTCTTGTCTATATTTCTGAAAACTACATA 4748
 Db 14755 TAGGTACTGAACAGATCAAGACTTTCTTTC-TGTCTATATTTCTTAAACATACATA 14813
 Qy 4749 TAAAGAACTTATATAGCATTTTGTGCTATTTCTTAAATA-CTTAAATGAT 4807
 Db 14814 TAAACATTTATTTACATGACCTTACATTTTAAAGGATTTTAAAGTATCTAGAAATAT 14873
 Qy 4808 TTAATGTATCTGGAGAAAGTGCATAGATATATCAATACCA-----TATATAG 4859
 Db 14874 TTAAGTATTAAGAGAAAGTTCAATAGTATATTTCAAAATCTACATATTTCTATAG 14933

Qy 4860 GAAATGACATCTGCAGATTTTGTCTGTGCTGAGGTTCTGGAAGAAATCCCTGTAA 4919
 Db 14934 GAAATGACATCTGCAGATTTTGTGTGTGTCAGAGGTTCTGGAAGAAATCCCTGTG 14993
 Qy 4920 TACACAAAATGACATCTTCGATGTGAATCTGAAGTCCCAAGCATCTATACATGCA 4979
 Db 14994 TACTGAGAGTACATCTTATTAAGACCTGAGCTGAGGCTCCAAAGCATGACATTAAG 15053
 Qy 4980 ATTCCAAAATTTGCTGCTCCCACTTCTAGAGATTTGCCCTCATCTTGTATCTTCA 5039
 Db 15054 ACTCCAAAATTTGCTGCTCCCTATATGCTTGAAGATTCCTCTATTTGTGTCTTAC 15112
 Qy 5040 TGTCTCCAGGACATTAAGCATCTTCAAGTCTTATGAAAAAGACGAGGAGAGAGGC 5099
 Db 15113 TGCTGCTCCACACCTCAGCATTCCTCAGCTGATGG-AGAAAGAGAGGAGAGAGGC 15171
 Qy 5100 TTTGCTCTCTATTAATCCCATGAGCAGGACTTGTCTGTGACATTTGTGATCTTC 5159
 Db 15172 TGTCTCTCTTATTTATGATCCATGAGCAGGACTTGTCTTACATCATCTTCTTACTTC 15231
 Qy 5160 CACTTAAACAGCAGCTGCTCATGAGATGATTCAGCATCAAGAAACTGGAGTGTGGT 5219
 Db 15232 CATTCACAGAACTGTGATCATGATGATTCAGCATCCAGAAACTGGAATGTGTGT 15291
 Qy 5220 CTTGTGTGCTGTGATCTTCAAGAAAGTTATGTGACCAAAAAGAA-ATCTTG 5275
 Db 15292 CTTGTGTGCTGTGATCTTCAAGAAAGTTATGTGATCTGAAAGAAATGATATGG 15351
 Qy 5276 GGGACACAGAGTCTTCAAGCTTCAAGCTGTCTGATCTGTGCTCATCATGAAT 5335
 Db 15352 GGCATGAGAGTGGCTTCAAGCTTCAAGCTGTGATCTGATCTGATCTGATCAATCAAGT 15411
 Qy 5336 TTTCAAGAACTCTCAGAAATTAATTAATGATGAGGAGAGAACTGAGTCTGCTG 5395
 Db 15412 TTTCAAGAGTCTCTCAAAAATTAATGATGAGAGAGAAAGCTGAGAGTCTATG 15471
 Qy 5396 CAGACATCCAGAGACAGAGGCTGTGATTTGAGACTGCTCTTCAATGTTGTGAAC-CAGAA 5454
 Db 15472 CAGGCTCTCAGAGACAGAGGCTTATCTGATCTGCTCTTCAATGATGTATCTCAGAA 15531
 Qy 5455 AACCTTAAATCTCTAGGCTTGTGCTTCAATGATTAATGAGATTAATCATGAA 5514
 Db 15532 GATCTTTGATCTTTAAGCTCTCAGAGCTTGTGTCAATGAGAGATTAATCAATGA 15591
 Qy 5515 CAGCTTTAAAGAACATCATGATGATTTAAACAACATGCTAAATGTTGTATCCAG 5574
 Db 15592 TGTCTTTGAGAAC-----ACAGGCTAAATGCTGCAATCTACAG 15633
 Qy 5575 TGAAGAG-ACAGGATGATCTTACTCTTAAGATCTTGGGTTTATGAGAAAGCAA 5633
 Db 15634 TGAACAGAAACAGGATGATGATTTATTTATGAGCTTCAAGGTTTGTAAAGAAAGCAA 15693
 Qy 5634 CATATCATCATATCATATGATGAGCAAC-----AGTTATGCTCTGAGTGTGAT 5687
 Db 15694 TATATCATATCTTAAATAAGCGCAAACTTAAATCTTAAATGCTCTCAATGTGAT 15753
 Qy 5688 AACAGAGTCTCTCTTCTCTCCATTTCTCTTTTGGGCAATGAGAGCTGTGACGCT 5747
 Db 15754 AACAGAGTCTGTGTTTCTCTTCACTGAGGCTCTCAGCTCTGATCTCTTGTACTCCAG 15813
 Qy 5748 GTCTCTTAAAGAGCTCATGATGATGATGATCACTCTCTGATCTCTTACTCCAG 5807
 Db 15814 GTATCTCTAAGAGAGCTCAGAGATGATGATCACTCTCTGATCTCTTGTACTCCAG 15873
 Qy 5808 AGAGAGATGATCTTCTTCTTCACTGAGAGCTCTCCCAATGAGATTTCTTGGATTC 5867
 Db 15874 AGAGAGGTTGTCTTCTTCTTCACTGAGGCTCTCAGCTCTGATCTCTTGTACTTC 15933
 Qy 5868 AGACCAACGTTGCTCTGATGAGCAAAAAGAACTCTGATTTGTTCTTAAATGACAC 5927
 Db 15934 AGACCAACGAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 15993

QY 5928 GCACTACCTATT-TCCTTATTTGCTTCTCATCTGATCGAGTTGT 5986
 DB 15994 TCACCTACCTATTGTTGTTCCCTATTGATTGTTCTTCTCATCTGAGTTGT 16053
 QY 5987 TTAGTTA-ATTTTTTTTTTACGCCCAATTTGACGTGCACTTGATTTAACTGAG 6045
 DB 16054 TTAGTTAGCTTTTTTTTTTACGCCCAATTTGACGTGCACTTGATTTAACTGAG 16113
 QY 6046 AATGCTCTCTACTTTACCCCCCTTACATGATTAATGACATAGAGTGTGGT 6105
 DB 16114 AGTCATTTCTGCTT-CCCCCTCATCATATTAATGACATAGAGTGTGGT 16171
 QY 6106 CCAAGGGCTGTGAAAA-ATGATCATGAGTCAAGCCCTGCTGGGCTCATTTCACT 6164
 DB 16172 CCAAGGGCTGTGAAAAAGGTGACCTTGAGGAGCTGTGCTGGTCCAAATCCGACT 16231
 QY 6165 ATATAATATAAACCCTCCGCAAAATATATCTCTTTAATCTTAACTTCAATTC 6224
 DB 16232 GTGTAAAG-CCAGGGCAAAATTACTCACTTTTGTGCTTAATGACTTCTC 16284
 QY 6225 TGCAATACAGAAATATACTAATTTTACCTCTAGGCTTTCAGATGATTAAGAAGC 6284
 DB 16285 TGCCATATAGAAATATTAATTAATCTTACCTCTAGGCTTTCAGATGATTA-AGAC 16340
 QY 6285 AATACCTAATAAATGTCATATGCTGCTGTATTCCTCCAAATTAAGCTTAATCTCAT 6344
 DB 16341 AATCCCTAATAAATCTCATATGCTGCTGTATTCCTCCAAATTAAGCTTAATCTCAT 16400
 QY 6345 TCTCCAGTTGAAATTTGCAATGATATCTCTTTTAAACCAAGCCCTTACCTTCTCTA 6404
 DB 16401 TCACCAAGTTGAAATCTACATGAAATCTCTTTTAACTCCAAAGCCCTGAA-CTCTCTA 16457
 QY 6405 TTTCCAGTATGAGACCTCTCTATACAAATGTTGATGATCAAAAGAAAGCTTACCAAG 6464
 DB 16458 TTTCTCTATGAGACCTCTCTATACAAATGTTGATGATCAAAAGAAAGCTTACCAAG 16517
 QY 6465 ATCTCCGAAAGAGAAATGAAATAGTTTACATTTGATGATCTACAGCAAACTTATG 6524
 DB 16518 ATCTCTGATGAGAGAAATGAAATGTTTGAATGATGATCAAGCAAACTTATG 16577
 QY 6525 AGTCCCAATCATATTTCCCACTTCAATTTACCTGCTGAGTGGCACTGAG-CTCTACC 6583
 DB 16578 AGTCTCCCAAGAAATTTGCTACATTTCAATTTACCTGCTGAGGCTGAGCTGACC 16637
 QY 6584 CTCACCTACTCTTCTCTGTTCTATTTGCTGAGCAATTTGAGCTGAGCCCAAGCCCTACC 6643
 DB 16638 CAGGCTACTCTTCTTCTTCTATATTTGCTGAGCAATTTGAGCTGAG-CCACACCTTACC 16696
 QY 6644 CAACACTGTGTACAAA-TGCTTCTAGGGGTTCCGCAAGCACTGAGCTTATTT 6702
 DB 16697 CAACACTGTGTACAAAAGATTTCTAGGGGCTCAGAAAGCACTGAGCTTATTT 16756
 QY 6703 TAAAGGACATCATGCTGCAATTTCAAGT-TTTGGGCACTCATCATCTTCTTCAAC 6761
 DB 16757 CAAGGACATCATGCTGCAATTTCAAGTATTTGGGCACTCATCATGAAAGCTTCTTCAAC 16816
 QY 6762 ACAGATAGAGCTGTCCAAATTAAGATTTGATGAAATTTCTCTCTAATATA 6821
 DB 16817 ACAGAGGA-ACAAATTAATTTGATGAAATTTT-TACTCTCATATA 16868
 QY 6822 TGTGTGTGTTTAATGCTTACATTTGCTTCAATTTTATTTTCAATTTTCAATCAATC 6881
 DB 16869 TGTGTGTGTTTCAACCTTTAATTTGTTTATTTTATTTTCTTTTATTTCAAGC 16928
 QY 6882 TACCATTTGCAATGAGCTTCTCATGATGATTTCTTCAATTTGATGAAAGCTTTATGAAA 6941
 DB 16939 TATCATCTACATTTAGCTCTTCAATTTCAATGATTTTCACTGATGAAATTTTACAAA 16988
 QY 6942 GCACATGCTGCTTATGGA-ATAGGCACTAGAGATTAATTAATTAATTTGCTGCTG 7000
 DB 16989 ACACATGCTGCTTCACTAGAGCTGAGCACTGGGAAATATAA-ATTTT-GTGTCTCTC 17040
 QY 7001 TCTGCAATGACTG-ACACACTGAGTTATTTCTACCAACAGGCTCCGCAATTTTCA 7058

DB 17041 TCTGTATATACGGACACACACTGAGCAATTTCTAGCTACCAAGGCCAGTATCATCA 17100
 QY 7059 CAT--CTTAGGAAAGATCCATTTTCTCTGTTCTATATGATGATTTTCTCTGT 7115
 DB 17101 CTGTCTTAGGAAAGATCCATTTTCTCTGTTCTATATGATGATTTTCTCTGT 17160
 QY 7116 CCAGATGATGAGCTGCTGATGAGGCTGACAAACCACTTTTGTATGATTAAT 7175
 DB 17161 CCA-TGATGACCACTTTTATCATGAAAGTGTCC-CTTTATCCAACT 17209
 QY 7176 CTTCACTGAGAAATTTTCAAGCCAGATGATGATGATGATGATGATGATGATGATGAT 7235
 DB 17210 CTTCACTGAGAAATTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 17268
 QY 7236 GCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7295
 DB 17269 GCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17328
 QY 7296 TGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7355
 DB 17329 TTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17388
 QY 7356 AGGAC-CCGACCTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7414
 DB 17389 AGGACCTCTCACTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17448
 QY 7415 AATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7474
 DB 17449 AATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17507
 QY 7475 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7534
 DB 17508 GGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17567
 QY 7535 CAACATCCAGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7589
 DB 17568 CAACATCCAGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17627
 QY 7590 ----TCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7645
 DB 17628 TTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17687
 QY 7646 TGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7703
 DB 17688 TGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17747
 QY 7704 CGCTGTCTCATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7763
 DB 17748 CACCTTCTCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17804
 QY 7764 CATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7823
 DB 17805 GATTTTGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17864
 QY 7824 TCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7883
 DB 17865 CACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17924
 QY 7884 GCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7943
 DB 17925 GCACTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 17984
 QY 7944 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8003
 DB 17985 AGGATG-ATGCTGCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 18042
 QY 8004 GTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8063
 DB 18043 GTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 18102
 QY 8064 TACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8123

Db 18103 CACAGTCACATGAGTATTTTATGATGATGTTCTCTGAGGGGTCACAGCCGGCTCTGCT 18162

Qy 8124 GGTGAGATTTCTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 8193

Db 18163 GGTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 18222

Qy 8184 CTTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 8243

Db 18223 GGTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 18282

Qy 8244 TTTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 8303

Db 18283 TTTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 18328

Qy 8304 CTTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 8363

Db 18329 CTTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 18388

Qy 8364 GGTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 8423

Db 18389 GGTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 18448

Qy 8424 TTTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 8483

Db 18449 TTTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 18508

Qy 8484 TTTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 8539

Db 18509 TTTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 18568

Qy 8540 TTTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 8599

Db 18569 TTTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 18627

Qy 8600 TTTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 8618

Db 18628 TTTGAGATCTTGTGATCCCGAAGATCCGCTGACAGGCTGTACGTACATCTT 18646

RESULT 10
AC107948/1 156839 bp DNA linear PRI 19-JUN-2002
LOCUS Homo sapiens chromosome 11, clone RP11-81D23, complete sequence.
DEFINITION AC107948
ACCESSION AC107948
VERSION AC107948.7 GI:21909529
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 156839)
Birtten,B., Nusbbaum,C. and Lander,E.
Homo sapiens chromosome 11, clone RP11-81D23
Unpublished
2 (bases 1 to 156839)
Birtten,B., Linton,L., Nusbbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,J.,
Choepl,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gaidyna,S.,
Gardyna,S., Goid,S., Goyette,M., Hume,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Laroque,K., Lamazares,R.,
Lander,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,J., Marguis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrim,J., Menus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
Peterson,K., Phunkhang,P., Piere,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

Straus,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (24-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 156839)
Birtten,B., Linton,L., Nusbbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,J.,
Choepl,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gaidyna,S., Goid,S., Goyette,M., Hume,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Laroque,K.,
Lamazares,R., Lander,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Marguis,N.,
Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,
Mlenga,V., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Piere,N.,
Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C.,
Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,
Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S.,
Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 156839)
Birtten,B., Nusbbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
Camarata,J., Chang,J., Chazaro,J., Choepl,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gaidyna,S., Goid,S., Goyette,M., Hume,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lander,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Menus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Piere,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (19-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 19, 2002 this sequence version replaced gi:21307060.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L2483
Center clone name: 81_D_23
----- Location/Qualifiers
1. 156839
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

/chromosome="11"
/map="11"
/clone="RP1-81D23"
/clone_11b="RP1-11 Human Male BAC"
repeat_region
complement(1..80)
/rpt_family="L2"
complement(484..779)
/rpt_family="AluSx"
complement(794..1094)
/rpt_family="AluSx"
complement(1100..1396)
/rpt_family="AluSg"
1435..1692
/rpt_family="AluJo"
complement(1711..1939)
/rpt_family="MIR"
complement(2245..2344)
/rpt_family="AluSg"
2346..2384
/rpt_family="AluSg"
complement(2791..3083)
/rpt_family="AluSx"
3690..3863
/rpt_family="L2"
5060..5270
/rpt_family="AluSg"
5900..6027
/rpt_family="PRAM"
6504..6562
/rpt_family="CCCG"
7387..7502
/rpt_family="L2"
7738..7826
/rpt_family="MIR"
complement(8164..8273)
/rpt_family="LMC4"
8379..9017
/rpt_family="L3"
10028..10067
/rpt_family="(TG)n"
10386..10494
/rpt_family="L2"
11713..11767
/rpt_family="L2R56"
11768..12065
/rpt_family="AluSx"
12066..12116
/rpt_family="L2R56"
complement(12117..12389)
/rpt_family="AluY"
12390..12706
/rpt_family="L2R56"
complement(12901..13000)
/rpt_family="L2R49"
complement(13001..13348)
/rpt_family="AluJo"
complement(13350..13442)
/rpt_family="L2R48B"
complement(13460..13680)
/rpt_family="L2R48"
complement(13681..13989)
/rpt_family="AluSg"
complement(13990..14172)
/rpt_family="L2R48"
complement(14305..14325)
/rpt_family="L2R23"
14326..14634
/rpt_family="AluJb"
complement(14635..14757)
/rpt_family="L2R23"
complement(14759..15047)
/rpt_family="AluSx"
complement(15090..15395)

```

```

/rpt_family="AluJo"
complement(15396..15689)
/rpt_family="AluSx"
repeat_region
complement(15719..15812)
/rpt_family="PRAM_C"
complement(15814..16103)
/rpt_family="L2R23"
17515..17811
/rpt_family="AluSx"
complement(18003..18119)
/rpt_family="MER81"
complement(19303..19604)
/rpt_family="AluJb"
complement(19814..19962)
/rpt_family="PRAM"
complement(20403..20461)
/rpt_family="L3"
complement(20478..20790)
/rpt_family="AluJb"
21558..21593
/rpt_family="(T)n"
complement(22105..22314)
repeat_region

Query Match 17.8%; Score 1535.8; DB 9; Length 156839;
Best Local Similarity 91.2%; Pred.No. 0;
Matches 1699; Conservative 0; Mismatches 152; Indels 12; Gaps 6;

QY 6750 ATTCTTCTCAACAGATGAGCTGTCCAAATGAAATTCGATGAAATGAAATTTCTT 6809
DB 156837 ATCTCTCTCAACAGATGAGCTGTCCAAATGAAATTCGATGAAATTTCTT 156778

QY 6810 CATCTAATTATAGTGTGTCTCTATGCTTACATGATGATGATGATGATGATGATGAT 6869
DB 156777 CATCTAATTATAGTGTGTCTCTATGCTTACATGATGATGATGATGATGATGAT 156718

QY 6870 TTCATCAATTCACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 6929
DB 156717 TTCATCAATTCACATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 156661

QY 6930 CGTTATGAAAGACATGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 6988
DB 156660 TGTATATGAAAGACATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 156601

QY 6989 AATGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7048
DB 156600 AATGATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 156541

QY 7049 CATTTTCACAC--ATCTAGCGAAGATCCATTTTCTCTGTGTGTGTGTGTGTGTGTGT 7106
DB 156540 CATTTTCACACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCAT 156481

QY 7107 TTTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7166
DB 156480 CCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 156421

QY 7167 ATCTGAATTCCTCCACCTGAGGAAATTTCCAGGCGCAGATGATGATGATGATGATGATG 7226
DB 156420 ATCTGAATTCCTCCACCTGAGGAAATTTCCAGGCGCAGATGATGATGATGATGAT 156361

QY 7227 ACAGCACTGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7286
DB 156360 AAAGCCCTGCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 156301

QY 7287 GAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7346
DB 156300 GAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 156241

QY 7347 CTGGATTTGAGGACCCCACTTTTGTATGATGATGATGATGATGATGATGATGATGATG 7405
DB 156240 CTGGATTTGAGGACCCCACTTTTGTATGATGATGATGATGATGATGATGATGATG 156182

QY 7406 CTCCTCTTTAAATGAGACAGTAAATCCACATGCGAGGATGATGATGATGATGATGAT 7465
DB 156182 CTCCTCTTTAAATGAGACAGTAAATCCACATGCGAGGATGATGATGATGATGATGATG 156123

```


Db 156181 CTCTTTTAAATGAGACAGTAATCCATACGAGGCTGCTGGGAGATACAGAT 156122
 Qy 7466 CAACACCTGGTATCAATCTGGTTTCTGTTTCAAGGCTCATGAGCTGGGTTTCTGA 7525
 Db 156121 GATACAGCTGGTATCATCTGGTTTGTGTTCCAGGGGACAGAGCTAGGTTTCTGA 156062
 Qy 7526 GCATGATTCACCACTCCAGCTTTGGGTACAGAACTGACCAATCAACGAGCTGAGG 7585
 Db 156061 GCATGATTCACCACTCCAGCTTTGGGTACAGAACTGACCAATCAACGAGCTGAGG 156002
 Qy 7586 AGACTCTTGTATACAGAGACCTGAGCTTCAAGGGGCTGACGATGATGGTTCCCTTG 7645
 Db 156001 AGACTCTTGTATACAGAGACCTGAGCTTCAAGGGGCTGACGATGATGGTTCCCTTG 155942
 Qy 7646 TCGGCTGACAGAAAGCGGCTTGTGCTGTGGCTCTGGGCTGCGGATGCGAGAAAG 7705
 Db 155941 TCGGCTGACAGAAAGCGGCTTGTGCTGTGGCTCTGGGCTGCGGATGCGAGAAAG 155882
 Qy 7706 CTGCTCATCTTACATCTCTCAACCTGGTCCGCGGCACTTCTTCTTACGGGCA 7765
 Db 155881 CTGCTCATCTTACATCTCTCAACCTGGTCCGCGGCACTTCTTCTTACGGGCA 155822
 Qy 7766 TTATATGTTGCGCGCTTACGCTTCAATATCCGCAATCCATGCCAAATCCCACTG 7825
 Db 155821 TTATATGTTGCGCGCTTACGCTTCAATATCCGCAATCCATGCCAAATCCCTG 155762
 Qy 7826 CTGATGACCTTTCCTTCTTATAGGCTTACGCTGAGGCTGAGGCTGAGCAACGAGC 7885
 Db 155761 CTGATGACCTTTCCTTCTTATAGGCTTACGCTGAGGCTGAGGCTGAGCAACGAGC 155702
 Qy 7886 GCTGCTGCTCATCTGTTGGGCTTCTGTGACCACTGCGCGGCTGAGATCTGCTAT 7945
 Db 155701 GCTGCTGCTCATCTGTTGGGCTTCTGTGACCACTGCGCGGCTGAGATCTGCTAT 155642
 Qy 7946 CGGCTGATGATCTGCTGCTGCGGCTGCTGCGGCTGCGGCTGCGGCTGCGGCTG 8005
 Db 155641 CGGCTGATGATCTGCTGCTGCGGCTGCTGCGGCTGCGGCTGCGGCTGCGGCTG 155582
 Qy 8006 TCTGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8065
 Db 155581 TCTGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 155522
 Qy 8066 CAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8125
 Db 155521 CAATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 155462
 Qy 8126 TCAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8185
 Db 155461 TCAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 155402
 Qy 8186 TCAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8245
 Db 155401 TCAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 155342
 Qy 8246 CCAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8305
 Db 155341 CCAGATGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 155282
 Qy 8306 TGTGCGCTTAAAG 8365
 Db 155281 TGTGCGCTTAAAG 155222
 Qy 8366 AGCGTCAAAATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8425
 Db 155221 AGCGTCAAAATAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 155162
 Qy 8426 AGGTGATGAG 8485
 Db 155161 AGGTGATGAG 155102
 Qy 8486 TGAG 8541
 Db 155101 TGAG 155042

Qy 8542 CCTGCAACCTTGAACAATATATGATTTTCTTACCTTCTGCTTCAGAAATGCTCA 8601
 Db 155041 TCCTGCAACCTTGAACAATATATGATTTTCTTACCTTCTGCTTCAGAAATGCTCA 154982
 Qy 8602 GCG 8604
 Db 154981 GTG 154979

RESULT 11
 AC079120/c
 LOCUS AC079120 104547 bp DNA linear PRI 09-JAN-2002
 DEFINITION Homo sapiens BAC clone Rpl1-345M24 from 2, complete sequence.
 AC079120
 VERSION AC079120.6 GI:15638822
 KEYWORDS HMG.
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 104547)
 Sultison,J.B. and Waterston,R.
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)
 99063792
 MEDLINE
 PUBMED
 9847074
 2 (bases 1 to 104547)
 Kalicki,J., Cotton,M. and Elliott,G.
 The sequence of Homo sapiens BAC clone Rpl1-345M24
 Unpublished (2001)
 3 (bases 1 to 104547)
 Waterston,R.H.
 Direct Submission
 Submitted (18-AUG-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 4 (bases 1 to 104547)
 Waterston,R.H.
 Direct Submission
 Submitted (18-SEP-2001) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 5 (bases 1 to 104547)
 Waterston,R.H.
 Direct Submission
 Submitted (09-JAN-2002) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 On Sep 18, 2001 this sequence version replaced gi:14488391.

REFERENCE
 AUTHORS
 JOURNAL
 TITLE
 JOURNAL
 COMMENT

Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: sapiens@wustl.edu
 Summary Statistics
 Center project name: H_NH0345M24

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 Mapping information for this clone was provided by Dr. John D.

Db 85598 TCTATGAGCAGATCTCCCTAATTAACCAAAACGAGGAAGCATTAATAAAAAAGA 85539
 Qy 58 AACTATGACCACTGATGATTAATATACATGACAGAAATCCCCAAATACTAGCT 117
 Db 85538 AACTACAGACCAATATCCCTGATGAACATAGATCAAAAGTCTCCGCAAAATACAGCT 85479
 Qy 118 AACCATTCGACGATATCAAGAATAATCCACCTTGTCAAGTGGTTTCATACCA 177
 Db 85478 AACCAATTCGACGATATCAAGAATAATCCACCTTGTCAAGTGGTTTCATACCA 85419
 Qy 178 GGGGTGCA-GGATAGTTAATCATACAGAGTCAATTAATGATATCATCATAAACGA 236
 Db 85418 GGGATGAGGATGATTTATATCTCAAGTCAATTAATGATATCATCATAAACGA 85359
 Qy 237 ATTAATAACCAAAATCATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAATC 296
 Db 85358 ATTAATAACCAAAATCATGATCATCTCAATAGATGAGAAAAACATTTGACAAATC 85299
 Qy 297 TAAATTTCTTTAATTAATTAACCTTCAGCAAAATCCACATAGAAAGACATTAAT 356
 Db 85298 CAGATGCTTTAATTAATTAACCTTCAGCAAAATTCGACAGAAAGGACATTAATAG 85239
 Qy 357 GTATTAATAAGCATATATGACGACCAAGCAAACTTATTAATGAGGAAAAAGTTG 416
 Db 85238 GTATTAATAAGCATATATGACCAAGCAAACTTATTAATGAGGAAAAAGTTG 85179
 Qy 417 AAAACATTTGCTGAGAACTGAAACAAGAAAGATG-CTACTTTCACCACTTCTATTC 475
 Db 85178 AAAACATTTGCTGAGAACTGAAACAAGAAAGATG-CTACTTTCACCACTTCTATTC 85119
 Qy 476 AACATTAATGAGAAAGTTTATGCAAGCAATTCAGCAAGAAAGAAATAGAGGAC 535
 Db 85118 AACATTAATGAGAAAGTTTATGCAAGCAATTCAGCAAGAAAGAAATAGAGGAC 85059
 Qy 536 CAATTAATTAAGAAAGTCAATCTGCTGTTCACTGATGATGATTAATTAATTA 595
 Db 85058 CAATTAATTAAGAAAGTCAATCTGCTGTTGCTGTTGCTGATGATGATTAATTA 84999
 Qy 596 GAAACCTTAAGAACTGATCAAGAAAGCTCTGATGATGATTAATTAATTAATTA 655
 Db 84998 GAAACCTTAAGAACTGATCAAGAAAGCTCTGATGATGATTAATTAATTAATTA 84939
 Qy 656 TCGAGATTAAGAACTGATCAAGAAAGCTCTGATGATGATTAATTAATTAATTA 715
 Db 84938 TCGAGATTAAGAACTGATCAAGAAAGCTCTGATGATGATTAATTAATTAATTA 84879
 Qy 716 GCTGAGATTAAGAACTGATCAAGAAAGCTCTGATGATGATTAATTAATTAATTA 769
 Db 84878 GCTGAGATTAAGAACTGATCAAGAAAGCTCTGATGATGATTAATTAATTAATTA 84819
 Qy 770 TTAAGATTAATTAAGAAAGTGAAGAAAGCTCTGATGATGATTAATTAATTAATTA 829
 Db 84818 TTAAGATTAATTAAGAAAGTGAAGAAAGCTCTGATGATGATTAATTAATTAATTA 84759
 Qy 830 CT----GACATCATGATGACAAAGATGAAAGCAATCCATGCTCATGATGAT 885
 Db 84758 CTGTTAATAATTAAGATGACAAAGATGAAAGCAATCCATGCTCATGATGATGAT 84699
 Qy 886 AGATCAATTAATTAAGAAAGTGAAGAAAGCTCTGATGATGATTAATTAATTAATTA 945
 Db 84698 AGATCAATTAATTAAGAAAGTGAAGAAAGCTCTGATGATGATTAATTAATTAATTA 84639
 Qy 946 CCCACCAAAATATCATCATCTTTCAGAGAACTGATGATGATTAATTAATTAATTA 1005
 Db 84638 TCCATCAAAATTAATCATCATCTTTCAGAGAACTGATGATGATTAATTAATTAATTA 84580
 Qy 1006 ATGGAACAAAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1065
 Db 84579 GTGAGAC-----CAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 84533
 Qy 1066 AAGAAACAAATCTGAGGATCATCACTTCACTTCACTTCACTTCACTTCACTTCACT 1125
 Db 84532 AAGAAACAAATCTGAGGATCATCACTTCACTTCACTTCACTTCACTTCACTTCACT 84473

Qy 1126 ACCAAAAATCATGAGCACTGACATTAATACTAGCAATAGCAATAGCAATAGCAATAG 1185
 Db 84472 ACCAAAAATCATGAGCACTGAGTTTAAATATAGAAATATAGCAATAGCAATAGCAAT 84413
 Qy 1186 AATCCAAATTAAGCCAAATTAATATAGCAATGATTTTGAACAAAGCAAAACAAAT 1245
 Db 84412 AATCCAAATTAAGCCAAATTAATATAGCAATGATTTTGAACAAAGCAAAACAAAT 84353
 Qy 1246 ATAAAGTGGGAAAGAAAGCAATTTAGTTAACAATAGTGTGATATTTGGCAAGCA 1305
 Db 84352 ATAAAGTGGGAAAGAAAGCAATTTAGTTAACAATAGTGTGATATTTGGCAAGCA 84293
 Qy 1306 TGTGAAGAAATGAAGCTGATCCCTGTTCTCACTTAATTAACAATAATGATCAAG 1365
 Db 84292 TGTGAAGAAATGAAGCTGATCCCTGTTCTCACTTAATTAACAATAATGATCAAG 84234
 Qy 1366 GATCCAAAGCTTAATTTGAGACCTTAACCAATTAATTTGAGAAATATATCAAGAA 1425
 Db 84233 AATCCAAAGCTTAATTTGAGACCTTAACCAATTAATTTGAGAAATATATCAAGAA 84174
 Qy 1426 AATGCTTCTAGACATTCATTAAGCAAGAACTTATGAGCAAGAAAGCAAGAAAGT 1485
 Db 84173 AATGCTTCTAGACATTCATTAAGCAAGAACTTATGAGCAAGAAAGCAAGAAAGT 84114
 Qy 1486 AACAAAAACAAATTAATTAATGATGAGCTTAATTAATTAATTAATTTGCGAGCA 1545
 Db 84113 AACAAAAACAAATTAATTAATGATGAGCTTAATTAATTAATTAATTTGCGAGCA 84054
 Qy 1546 AACATCAATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1605
 Db 84053 AACATCAATTAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 83995
 Qy 1606 AGCATGCACTAAGCACTAATTAATCCGATTCACAAAGAACTGAAACAAATTCAGCA 1665
 Db 83994 TATTTCTGACAAAGATTAATTAATGAGAACTTACAGAAATTCAGCAATTCAGCA 83935
 Qy 1666 GAAAGCAAAATCCATGAAAGAGTGGGCTAAGCAATGAATTAAGCAATTTCTCA 1725
 Db 83934 GAAAGCAAAATCCATGAAAGAGTGGGCTAAGCAATGAATTAAGCAATTTCTCA 83875
 Qy 1726 AGATTAATCAATTTGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1784
 Db 83874 AGATTAATCAATTTGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 83815
 Qy 1785 ATGTAATCAAGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 1844
 Db 83814 ATGTAATCAAGCACTGATGATGATGATGATGATGATGATGATGATGATGAT 83755
 Qy 1845 AATCTTAATAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1904
 Db 83754 AATCTTAATAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 83695
 Qy 1905 TGTGGAATGATTAATTTGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1964
 Db 83694 TGTGGAATGATTAATTTGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 83635
 Qy 1965 CTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2001
 Db 83634 CTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 83575
 Qy 2002 -----TAAATAT 2008
 Db 83574 AAGTATGATTAATGCAAAAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 83515
 Qy 2009 GTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2040
 Db 83514 GTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 83455
 Qy 2041 ACAATTCAGGCAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2100
 Db 83454 ACAATTCAGGCAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 83395

QY 2101 AGACCTATTATCTAAGTGGGTAAGTCTAGAGTAAGAAACCAAAATCATATGTTCTCAC 2160
 DB 83394 GGACCATATTTTAAAGTGAATTAACAGAGAGTGAAACCAAAATCATATGTTCTCAC 83335
 QY 2161 TTACAAGTGGGGCTTAAGTGTGAGACAGAAAGCATATGATATATGAACTCTGG 2220
 DB 83334 TCATTAAGTGGAGCTTAAGTATAGAGACAAAGGAT--AAGATATCTATGAGCTTTGG 83277
 QY 2221 GGACTTGAAGGGAAGATGGAAGAGAGGAGGATTAAGATCAACAATGGTACAT 2280
 DB 83276 GGA-TTGAGGGGAATGAGTGGAGAGGGGTGAGGATTAAGATTAACAATGGTACAT 83218
 QY 2281 GTACACTGCTCAGTGTAGTGGTGCACCAAAATCTCAAAATTAACAATAAGAACTTATC 2340
 DB 83217 GTACACTGCTCGGTGTGTAGTGGTGCACCAAAATCTCAAAATTAACAATAAGAACTTATC 83158
 QY 2341 CATGGAAGCAACACCACTG-TTCCCAAAATCCCAATGAATTAATTAATTAATTAAT 2399
 DB 83157 CATGGAAGCAACATCATCTGCTTCCCAAAATTAATTAATTAATTAATTAATTAAT 83098
 QY 2400 AATGATTTAATTTCAAGATTTAAAGTTCA 2433
 DB 83097 CAATGGCCAGAAACATATGAATAAGTTCA 83064

RESULT 12
 AC074016/c 179216 bp DNA linear HTG 09-AUG-2001
 LOCUS Homo sapiens chromosome 2 clone RP11-155P18, WORKING DRAFT
 DEFINITION
 AC074016
 AC074016.5 GI:15144340
 VERSION HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 179216)
 Waterston, R.H.
 The sequence of Homo sapiens clone
 Unpublished
 2 (bases 1 to 179216)
 Waterston, R.H.
 Direct Submission
 Submitted (10-JUN-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 On Aug 9, 2001 this sequence version replaced gi:9958173.

----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Project Information -----
 Center project name: H.NH0155P18

 Summary Statistics -----
 Sequencing vector: M13, 488
 Sequencing vector: plasmid, 448
 Chemistry: Dye-terminator Big Dye; 448 of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 178616 bases at least Q40
 Consensus quality: 178729 bases at least Q30
 Consensus quality: 178811 bases at least Q20
 Insert size: 174000; agarose-fp
 Insert size: 181802; sum-of-fragments
 Quality coverage: 15.53 in Q20 bases; agarose-fp
 Quality coverage: 14.71 in Q20 bases; sum-of-fragments

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is

FEATURES
 source
 1..179216
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 /clone="RP11-155P18"
 misc_feature
 1..82838
 /note="assembly_name:Contig10
 clone_end:17
 vector_side:left"
 misc_feature
 82839..179216
 /note="assembly_name:Contig11
 clone_end:82838
 vector_side:left"
 ORIGIN
 Query Match 17.6%; Score 1521.2; DB 2; Length 179216;
 Best Local Similarity 81.0%; Pred. No. 0;
 Matches 1983; Conservative 0; Mismatches 383; Indels 82; Gaps 15;
 QY 1 TGTATGAGCCCAATGCTACTTTAATGCCAAACGAGAAAGATTA--CAAAAAGAAA 58
 DB 50441 TGTATGAGCCCAATGCTACTTTAATGCCAAACGAGAAAGATTA--CAAAAAGAAA 50382
 QY 59 ACTATGACACGATACACGATGATATATATGATGATGATGATGATGATGATGATGATGAT 118
 DB 50381 ACTATGACACGATACACGATGATATATATGATGATGATGATGATGATGATGATGATGAT 50322
 QY 119 ACCCAATCCAGACGATATCAAGAAAGATATATATGATGATGATGATGATGATGATGATGATGAT 178
 DB 50321 ACTCAATCCAGACGATATCAAGAAAGATATATATGATGATGATGATGATGATGATGATGATGAT 50262
 QY 179 GGGTGCAGATAGGTTAATACATACAGATGATGATGATGATGATGATGATGATGATGATGAT 238
 DB 50261 GGGTGCAGATAGGTTAATACATACAGATGATGATGATGATGATGATGATGATGATGATGAT 50206
 QY 239 TAAAAACAAATATCAT 298
 DB 50205 TAAAAACAAATATCAT 50146
 QY 299 ACATTTCTTTATGATTAATTAATCTTGCAGAAATGCAATAGAAAGACATACCTTAATGT 358
 DB 50145 GCA-TGCTTTATGATTAATTAATCTTGCAGAAATGCAATAGAAAGACATACCTTAATGT 50087
 QY 359 AATTAAGCCATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 418
 DB 50086 AATTAAGCCATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 50027
 QY 419 AATATGTCCTGAGAACTGGAACAGAAAGATG-CTACTTACCACTTCTTAATCA 477
 DB 50026 AGATTTCCCACTGAGAACTGGAACAGAAAGATG-CTACTTACCACTTCTTCA 49967
 QY 478 CATATGATGAGAAATTTTGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 537
 DB 49966 CATATGATGAGAAATTTTGAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 49907
 QY 538 AATCAATTAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 597
 DB 49906 AATCAATTAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 49847
 QY 598 AATCAATTAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 657
 DB 49846 AATCAATTAAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 49787

658 AGGATCAAACTAAATCTAGACAAATGAGTACCTGTATACCAACATGACCAAGC 717
 49786 CAGATCAAAATTAATGTGACAAATCAGTACCTGTCTATACCTGACGCGACTAGG 49727
 718 TGGAATCAATCAAGAACTCAACCTTTTACATAGCTGTAAATAAAATCTTAA---- 773
 49726 TGGAATCAATCAAGAACTCAACCTTTTACATAGCTGTAAATAAAATCTTAA---- 49667
 774 ---GAATATTTCTTACCAAGAGGTGAAGACCTCTACAGAAATACTACAAACACAGC 830
 49666 TTGACATATACCTAACAGAGGTGTAGACCTGTACAGAAATACTTAAACACTGC 49607
 831 T---GAATCAATCAAGAACTCAACAGTGGAAACATCCGATGCTCAAGATGGGTAG 887
 49606 TAAAGAAATCAATATACCAACCAATGGAAACATCCGATGCTCAAGATGGGTAG 49547
 888 AATCAATATTTGTGAATAATGACATATTTGCCAAAGCAATCTACAGTTCAATGCAATCC 947
 49546 AATCAATATTTGTGAATAATGACATATTTGCCAAAGCAATCTACAGTTCAATGCAATCC 49487
 948 CACCAAAATATCATCATCTTCTTACAGAACTGAAATAAACTCTAAATTCATAT 1007
 49486 CATCAAAATATCACCATCTGTTTTCACAGAACTGAAATAAACTCTTAAATTCATAT 49427
 1008 GGAACACCAACCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 1066
 49426 GAACCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 49367
 1067 -----AGAACAAATCTGAGGAGCTCATTTACCA 1096
 49366 AAGCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 49307
 1097 TCTTCAATATATCTACAGAGCTTAAATCACCAAAATCAAAATCAAAATCAAAATCAAA 1156
 49306 ATTTCAATATATATATATATGAGCTACCAAAATCAAAATCAAAATCAAAATCAAA 49247
 1157 GGCAATATATCAATATGAAAGAGAAATCCAGAAATTAACCAATTAATTAATGAGC 1216
 49246 GGCAATATATCAATATGAAAGAGAAATCCAGAAATTAACCAATTAATTAATGAGC 49187
 1217 AACTGATTTTGAACAAAGCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 1276
 49186 AACTGATTTTGAACAAAGCAAAATCAAAATCAAAATCAAAATCAAAATCAAAATCAAA 49127
 1277 AATGATCTGAGATTAATGAGCAACCAATGAGAAATGAACTGATCCCTGCTC 1336
 49126 AATGATCTGAGATTAATGAGCAACCAATGAGAAATGAACTGATCCCTGCTC 49068
 1337 TCATTAATATCAAAATTTGATACAAAGTGAATCAAAATCAAAATTTGAGACTTAAAC 1396
 49067 TCATTAATATCAAAATTTGATACAAAGTGAATCAAAATTTGAGACTTAAAC 49008
 1397 ATAAATATTTCTAGAAAGTAAATCAAGAAATGCTTCTAGACATTCCTTGGGAAAGAC 1456
 49007 ATAAATATTTCTAGAAAGTAAATCAAGAAATGCTTCTAGACATTCCTTGGGAAAGAT 48948
 1457 TTCAATGAGCAAGAACCAAAATGAAATGCAAAATCAAAATTAATTAATGATAGACTTA 1516
 48947 TTCAATGAGCAAGAACCAAAATGAAATGCAAAATTAATTAATGATAGACTTA 48888
 1517 ATTAATATTAAGAGCTTTTGGGCAACAAATCAATTAAGAGCAAGCAAGCAAGCAAGC 1576
 48887 ATTAATATTAAGAGCTTTTGGGCAACAAATCAATTAAGAGCAAGCAAGCAAGCAAGC 48828
 1577 ACCGATGAGAGAAATCTTCAAAATCTGATCTGATTAAGAGCAAGCAAGCAAGCAAGC 1636
 48827 ACAGAGTGGAGAGAAATCTTCAAAATCTGATCTGATTAAGAGCAAGCAAGCAAGCAAGC 48768
 1637 CACAAAGAACTCAAAATCAAGCAAGAGAAAGCAAAATCCCATGAAAGAGTGGGCT 1696
 48767 TACAACAAATCAAAATCAAGCAAGAGAAAGCAAAATCCCATGAAAGAGTGGGCT 48710
 1697 AAGGACATGAAATCAAAATCTCAAAAGAGATATACAAATGGGCAACAAAC--AGGAA 1754

48709 AAGGACATGAAATGACAAATTTCTAAAGAGATTAACAAATTTGGCCAAACAAATATATA 48650
 1755 AAATGCTTAATCATCACTAATGATTAAGGAATGAAATCAACATGTAATGCAATACAC 1814
 48649 AAGTCTCAACATCACTAATGATTAAGGAATGAAATCAACAAATGATGATACAC 48590
 1815 CTATCTCTGCAAGATGCTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1874
 48589 CTATCTCTCAAGAAATGGCCATTAATCAAAATTAATTAATTAATTAATTAATTAATTAAT 48530
 1875 TGTGTGATTAAGGAACATTTTACCTGCTGGGGAATGTAATTAATTAATTAATTAATTA 1934
 48529 TGTGTGATTAAGGAACATTTTACCTGCTGGGGAATGTAATTAATTAATTAATTAATTA 48470
 1935 TGGAAACAGTGTGAATTTCTTAAGGAATTAAGGAATTAAGGAATTAAGGAATTAAGGA 1994
 48469 TGGAAACAGTGTGAATTTCTTAAGGAATTAAGGAATTAAGGAATTAAGGAATTAAGGA 48410
 1995 ATCCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2054
 48409 ATTC-----ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 48373
 2055 AAAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2114
 48372 AAAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 48316
 2115 AGTGGGTAATCACTAGAAATGAAACCAACATCATATGTTCTCACTTACAAAGTGGGGC 2174
 48315 AGTGAATTAATCAAGAAATGAAACCAACATCATATGTTCTCACTTACAAAGTGGGGC 48256
 2175 TAACTGTGAGACCAAGCAT--AGATGATTAATTAATTAATTAATTAATTAATTAATTA 2233
 48255 TAACTGTGAGACCAAGCAT--AGATGATTAATTAATTAATTAATTAATTAATTAATTA 48196
 2234 AGGA--TGGAAAGAGGAGGAGGATTAAGAACTACAAATGAGTACAGTGTAC 2292
 48195 AAGAGTGAAGAGGAGGAGGATTAAGAACTACAAATGAGTACAGTGTAC 48136
 2293 GGTATGGTGTGACCAAAATCTCAGAAATTAACCACTAAGAACTTACATGAAAGCAAA 2352
 48135 GGTATGGTGTGACCAAAATCTCAGAAATTAACCACTAAGAACTTACATGTAACCAA 48076
 2353 CACCACTGTTCCCAAAATCCCAATGAATTAATTAATTAATTAATTAATTAATTAATTA 2400
 48075 CACCTCTGTTCTTCTAATTAACCAATGAATTAATTAATTAATTAATTAATTAATTA 48028

RESULT 13
 AC007256/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AC007256 181150 bp DNA linear PRI 07-NOV-2001
 Homo sapiens BAC clone Rpl1-575C6 from 2, complete sequence.
 AC007256
 AC007256.5 GI:14327821
 HTG.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 181150)
 Sulston,J.E. and Waterston,R.
 Toward a complete human genome sequence
 Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE
 PUBMED
 9963792
 9847074
 2 (bases 1 to 181150)
 Du,F., Laplant,Y., Doeber,A. and Moore,B.
 The sequence of Homo sapiens BAC clone Rpl1-575C6
 Unpublished
 JOURNAL
 3 (bases 1 to 181150)
 Waterston,R.H.
 Direct Submission
 JOURNAL
 Submitted (07-APR-1999) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
4 (bases 1 to 181150)
Waterston, R.H.
Direct Submission
Submitted (19-APR-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
5 (bases 1 to 181150)
Waterston, R.H.
Direct Submission
Submitted (20-APR-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
6 (bases 1 to 181150)
Waterston, R.H.
Direct Submission
Submitted (07-JUN-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
7 (bases 1 to 181150)
Waterston, R.H.
Direct Submission
Submitted (03-JUL-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE
AUTHORS
TITLE
JOURNAL
8 (bases 1 to 181150)
Waterston, R.
Direct Submission
Submitted (07-NOV-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jun 7, 2001 this sequence version replaced g1:13677176.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.edu
----- Summary Statistics
Center project name: H_NH0575C06

COMMENT

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPc1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, R., Tateo, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACe3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-156B7, the clone sequenced to the right is RP11-53618, 200 bp overlap. Actual start of this

clone is at base position 1 of RP11-575C6; actual end is at base position 180956 of RP11-575C6.

The sequence from base position 134614 to 134714 is derived from PCR product of RP11-575C6 BAC DNA.

The clone RP11-575C6 contains a transposon in the vector.

FEATURES

source	Location/Qualifiers
1. 181150	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="2"
	/map="2"
	/clone="RP11-575C6"
	/clone_1ib="RPc1-11"
repeat_region	1. .354
	/rpt_family="L1"
misc_feature	91. .95
	/note="similar to Homo sapiens EST BG495251 (NID:g13456766)"
repeat_region	124. .153
	/rpt_family="(TG)n"
repeat_region	192. .228
	/rpt_family="(T)n"
repeat_region	330. .1055
	/rpt_family="L1"
repeat_region	553. .620
	/rpt_family="T-rich"
repeat_region	1056. .1363
	/rpt_family="Alu"
repeat_region	1210. .1243
	/rpt_family="(T)n"
misc_feature	1228. .1246
	/note="match to EST AW444732 (NID:g6986494)"
repeat_region	1364. .2264
	/rpt_family="L1"
repeat_region	3338. .3550
	/rpt_family="MIR"
repeat_region	3739. .3866
	/rpt_family="MIR"
misc_feature	3971. .4133
	/note="similar to Homo sapiens EST AU117593 (NID:g10932573)"
misc_feature	4041. .4495
	/note="similar to Homo sapiens EST A1499220 (NID:g4391202) to08f03.x1"
misc_feature	4053. .4639
	/note="similar to Mus musculus EST BF163999 (NID:g1044285)"
misc_feature	4092. .4830
	/note="similar to Homo sapiens EST AL528010 (NID:g12791503)"
misc_feature	4128. .4340
	/note="match to EST AV760501 (NID:g10918349)"
misc_feature	4184. .4742
	/note="match to EST A1052250 (NID:g3308241) oz21c07.x1"
misc_feature	4339. .4903
	/note="similar to Homo sapiens EST AU117593 (NID:g10932573)"
misc_feature	4485. .4491
	/note="similar to Homo sapiens EST BE788113 (NID:g10209311)"
misc_feature	4657. .5468
	/note="similar to Homo sapiens EST AL563032 (NID:g12912044)"
misc_feature	4712. .5400
	/note="match to EST AU131936 (NID:g10992290)"
misc_feature	4780. .5582
	/note="similar to Homo sapiens EST BG164297 (NID:g12671000)"
misc_feature	4963. .5682
	/note="similar to Homo sapiens EST AU124988

```
misc_feature (NTD:g10949704) "
5106..5633
/Note="similar to Mus musculus EST A1116325 (NTD:g3516649)
ue97c04.y1"
5159..5675
/Note="similar to Homo sapiens EST A1144381
(NTD:g11005902) "
5194..5761
/Note="similar to Homo sapiens EST BF529656
(NTD:g11617019) "
5237..6019
/Note="similar to Homo sapiens EST A1130508
(NTD:g10990862) "
5238..6043
/Note="similar to Mus musculus EST BF166846
(NTD:g11047198) "
5326..5653
/Note="similar to Homo sapiens EST BG390376
(NTD:g13283824) "
5361..6051
/Note="similar to Homo sapiens EST A1135897
(NTD:g10996436) "
5361..5812
/Note="similar to EST BG862638 (NTD:g14213176) "
5420..6048
/Note="similar to Homo sapiens EST BF788113
(NTD:g10209311) "
5551..6327
/Note="similar to Homo sapiens EST A1130771
(NTD:g10991125) "
5615..5984
misc_feature
17.6%; Score 1521.2; DB 9; Length 181150;
Best Local Similarity 81.0%; Pred. No. 0;
Matches 1983; Conservative 0; Mismatches 383; Indels 82; Gaps 15;
Query Match
1 TGTATGAAGCCATGTCTTAAATACCAAAACGAGAAAGATATA--CAAAAAGAA 58
Db 100770 TGTATGAAGCCATGTCTTAAATACCAAAACGAGAAAGATATA--CAAAAAGAA 100711
Qy 59 ACTATGACGAGTACGATGATATATACATGAGAAATCCCAACAAATCTAGCTA 118
Db 100710 ACTATGACGATATCTTGAATGACATGATGCAAAATCTTAAACAAATCTAGCTA 100651
Qy 119 ACCCAATCCACAGCATATCAAGAAATATCCATTTCAAGTGGTTTCAATACAG 178
Db 100650 ACTCAATCCACAGCATATCAAGAAATATCCATTTCAAGTGGTTTCAATACAG 100591
Qy 179 GGGTGCAGATAGTTAAATATCAATCAATTAATGTATCATCATCAATCAAT 238
Db 100590 GGGTGCAGATAGTTAAATATCAATCAATTAATGTATCATCATCAATCAAT 100535
Qy 239 TAAATAACAAAATACATGATCATCTCAATGATGCTGAAACATTTGCAAAATCTA 298
Db 100534 TAAATAACAAAATACATGATCATCTCAATGATGCTGAAACATTTGCAAAATCTA 100475
Qy 299 ACATTTCTTATGATTAACCTTCAGCAAAATGACATAGAAAGACATACCTTAATGT 358
Db 100474 GCA-TGCTTTATGATTAACCTTCAGCAAAATGACATAGAAAGACATACCTTAATGT 100416
Qy 359 AATTAAGCCATATATGACGACCAAGCAACATTAATCTGAATGGGAAAGTTGAA 418
Db 100415 AATTAAGCCATATATGACGACCAAGCAACATTAATCTGAATGGGAAAGTTGGA 100356
Qy 419 AACATTTCCCTGGAATGGAACAAGAGATG--CTACTTCAACACTTCAATCAA 477
Db 100355 AGCATTTCCCTGGAATGGAACAAGAGATG--CTACTTCAACACTTCAATCAA 100296
Qy 478 CATATGATGAGATTTTATGACGACCAAGCAATCAAGAGAAAGAAATCAAGGACCCA 537
Db 100295 CATATGATGAGATTTTATGACGACCAAGCAATCAAGAGAAAGAAATCAAGGACCCA 100236
Qy 538 AATCAATTAAGAGAGATCAAACTGTCTCTGTCACTGATGATATGATTGATTAACCTAGA 597
```

```
Db 100235 AATCAGTAAGAGAGAGATCAAACTATGTCTGTGTCGATGATATTAATGTATACCTACA 100176
Qy 598 AAACCTAAAGACTATCCAGAAAGCTCTTGAACGTATACATTAATTCAGTAAGTTTC 657
Db 100175 AAACCTAAAGATTCCTCCAAAAGCTCTTGAACGTATACATTAATTCAGTAAGTTTC 100116
Qy 658 AGATTAACAACTAAATGTACAACAAATGATAGCACTGCTATACCAACAGTGACCAAGC 717
Db 100115 CAGATCAAAATTAATGTGACAAATGATAGCTCTGCTATACCTACAGCGACTAGAG 100056
Qy 718 TGAATAATCAATCAAGAACTCAACACTTTTACAAATGCTGTAAAAAATTAATCTAA---- 773
Db 100055 TGAATAATCAATCAAGAACTCAACACTTTTACAAATGCTGTAAAAAATTAATGT 99996
Qy 774 ---GATATTTCTTACCAAGAGGTGAGAGACCTCTCAAGAGAAACCTACAAACACAGC 830
Db 99995 TTGACCTATACCTACCAAGAGGTGTAAGACCTGTACAGAGAAACCTATTAACACTGC 99936
Qy 831 T--GACATCATGATGACCAAAAGTGGAACACATCCCATGCTCATGATGGGTAG 887
Db 99935 TAAAGAAATCATATGATGACCAAAAGTGGAACACATCCCATGCTCATGATGGGTAG 99876
Qy 888 AATCAATATTTGTGAATAATGACATTTGCAAAAGCAATCTACAGTTCAATGCAATGCC 947
Db 99875 AATCAATATTTGTGAATAATGACATTTGCAAAAGCAATCTACAGTTCAATGCAATGCC 99816
Qy 948 CACCAAAATATCATCATCTTCTTACAGAACTGAAATAAAACAAATCTAAATTCATAT 1007
Db 99815 CATCAAAATATCATCATCTTCTTACAGAACTGAAATAAAACAAATCTAAATTCATATG 99756
Qy 1008 GGAACAAACAAACCAAAAAAAAAAAAAAAAACCCCGATAGCCAAAGCAAGACTTACCAAA- 1066
Db 99755 GAAACCAAAAGAAACCAACATAGCCAAAGCAAGACTTACCAAAAGCAAGCAAAAC 99696
Qy 1067 -----AGAACAAATCTGAGGATCATTAACCA 1096
Db 99695 AAGCAAAACAAAAATCAAAAAAGCAAAAAACCTTAACAAATCTGAGGATCATTAACCTG 99636
Qy 1097 TCTTCAAACTATATCTACAGGCTTATATCAACAAACATCATGCACTGACATTAACCTA 1156
Db 99635 ATTTCAACTATATATATATGCGCATGTCACCAAAACATCATGCTATCGATTAATATA 99576
Qy 1157 GGCACATAGACCAATGGAAGAAAGAGAAATCCAGAAATTAAGCCAAATATATATAGCC 1216
Db 99575 GGCACATAGACCAATGGAAGAAAGAGAAATCCAGAAATTAAGCCAAATATATATAGCC 99516
Qy 1217 AACTGATTTTGAACAAGCAAAACAAACATTAAGTGGGAAAGACATTTACTTAACA 1276
Db 99515 AACTGATTTTGAACAAGCAAAACCAATATTAAGTGGGAAAGACATTTACTTAACA 99456
Qy 1277 AATGCTGATGATTTTGGCAAGCCACATGTGGAAGATGAATCTGATCTCTGTCTC 1336
Db 99455 AATGCTGATGATTTTGGCAAGCCACATGTGGAAGATGAATCTGATCTCTGTCTC 99397
Qy 1337 TCACTTAATCAAAATTTGATACAGATGATCAAGACTTAATCTGAGACTTAACCC 1396
Db 99396 TCACTTAATCAAAATTTGATACAGATGATCAAGACTTAATCTGAGACTTAACCC 99337
Qy 1397 AATTAATTTCTAGAGATTAACATGAGAAATGCTTTTACACATTCCTTAGGCAAGAC 1456
Db 99336 AATTAATTTCTAGAGATTAACATGAGAAATGCTTTTACACATTCCTTAGGCAAGAT 99277
Qy 1457 TTCAATGCAAGAACCCAAAGTAAATGCAACAAACAAATTAATTAATGATGAGACTTA 1516
Db 99276 TTCAATGCAAGAACCCAAAGTAAATGCAACAAACAAATTAATTAATGATGAGACTTA 99217
Qy 1517 AATTAACATAAAGCTTTTGGCAAGCAAAACAAATCAATTAAGCAAGCAACCAACC 1576
Db 99216 AATTAACATAAAGCTTTTGGCAAGCAAAACCAATCAATTAAGCAAGCAACCAACC 99157
Qy 1577 ACCGATGAGAGAAATCTTCAACAACTGAAGCTTGAAGACTTAATTCGGGATC 1636
```


Db	99156	ACAGGCTGGAGAAATCTTCCATCATATCATCTGACAAAGACCTAATATCCAGATC	99097
Db	1637	CACAAGAACTCATAAATAATGACGAGAAAGAACAAATCCCATGAAAGAGTGGCT	1696
Db	99096	TACAAACAACTAAACAAATTTAGCAAGAAAAA--CAATATACCCTCAAAAATGGGCT	99039
Db	1697	AAGGATGATATGACATTTCTGAAAAGAAATATGAAATGGGCAACAA--AGGAAA	1754
Db	99038	AAGGATGATATGACATTTCTGAAAAGAAATATGAAATGGGCAACAA--AGGAAA	98979
Db	1755	AAATGCTTACATCATCTATATGATTTAGGAAATGTAATCAACACTGTATGCTATCCAC	1814
Db	98978	AAGGCTCAACATCATCTATATGATTTAGGAAATGTAATCAACACTGTATGCTATCCAC	98919
Db	1815	CTTACTCTGCAAGATGCTCATATTTAAATATCTAAATATATAGATGTTGGTGGCTC	1874
Db	98918	CTTACTCTGCAAGATGCTCATATTTAAATATCTAAATATATAGATGTTGGTGGCTC	98859
Db	1875	TGTGCTGATTAAGAAACCTTTACACGCTGTGGGAAATGTAATCTGGGCAACCTA	1934
Db	98858	TGTGCTGATTAAGAAACCTTTACACGCTGTGGGAAATGTAATCTGGGCAACCTA	98799
Db	1935	TGGAACAAAGTGTGGAAATTTCTTAAAGAACTTAAAGATGACATTTGATTCACGA	1994
Db	98798	TGGAACAAAGTGTGGAAATTTCTTAAAGAACTTAAAGATGACATTTGATTCACGA	98739
Db	1995	ATCCCATTAATATGAT	2054
Db	98738	ATTC-----AT	98702
Db	2055	AAAAAGATTAATATGAT	2114
Db	98701	AAAAAGATTAATATGAT	98645
Db	2115	AGTGGGGTAACTCAGGAATGGAACCAACATCATATGTTCTCACTTACAGTGGGGC	2174
Db	98644	AGTGAAGTAACTCAGGAATGGAACCAACATCATATGTTCTCACTTACAGTGGGAAC	98585
Db	2175	TAACTGTGAGACACGAGAGCAT-AGATGATATATATATGAACTCTGGGACCTTGAAGGGA	2233
Db	98584	TAACTGTGAGAGATCAAGACGATTAAGATTAACCAATGGAATTTGGGGACTCAGGGGGA	98525
Db	2234	AGGA-TGGAAGAGAGCGAGGATTAAGATCTACATATGGGTACAGTGTACACTGCTCA	2232
Db	98524	AAGAGTGGAAAGGGGTAGAGGATTAAGATCTACAAATTTGATGAGTGTACTGCTCG	98465
Db	2293	GGTATGGGTGACACCAAAATCTCAGAAATTTACCACTAAAGAACTTATCATGAGAGCAA	2352
Db	98464	GGTATGGGTGACACCAAAATCTCAGAAATTTACCACTTAAAGTACTTATCATGATTAACCAA	98405
Db	2353	CACCACTGTTTCCCAAAATCCCAATGAAATTAATATATATATATATATATATATATAT	2400
Db	98404	CACCTCTGTTTCTTAATTAACCAATGAAATTAATATATATATATATATATATATATAT	98357
RESULT 14	AC126613/c	125001 bp DNA linear PRI 30-JUL-2002	
LOCUS	AC126613	Homo sapiens 12 BAC Rpl11-715M6 (Roswell Park Cancer Institute Human	
DEFINITION	BAC library) complete sequence.		
ACCESSION	AC126613.6	GI:22003946	
VERSION	AC126613.6	GI:22003946	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	1 (bases 1 to 125001)		
AUTHORS	Munzly, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,		
	Albrooks, S.L., Amarantunge, H.C., Are, J.R., Ayale, M., Banks, T.,		
	Barbarta, J., Benton, J., Blinze, K., Blankenburg, K., Bonnin, D.,		
	Bouck, J., Bowler, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,		
	Bunay, C., Butch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,		
	Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chaver, D.,		
	Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhury, I., Christopoulos, C.,		
	Cleveland, C.D., Cox, C., Coyle, M.D., Dachtore, S.R., David, R.,		
	DeVella, M.L., Davis, C., Davy-Carroll, L., Dedrich, D.A.,		
	Delaney, K.R., Delgado, O., Dem, A.L., Ding, Y., Dinh, H.H.,		
	Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,		
	Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emeling, S.,		
	Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P.,		
	Frantz, P., Gabriel, A., Gao, J., Garcia, A., Garner, T., Garza, N.,		
	Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,		
	Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Hayak, P.,		
	Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M.,		
	Holloway, C., Hollins, B., Homel, F., Howard, S., Huber, J., Huliy, S.,		
	Hume, J., Ioshikawa, I., Jackson, L.E., Jacobson, B., Jia, Y.,		
	Johnson, R., Jolivet, S., Joudan, S., Karlsson, E., Kelly, S., Khan, U.,		
	King, L., Korah, J., Kovar, C., Kratovic, J., Kureishi, A., Landry, N.,		
	Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, J., Li, Z., Licharge, O.,		
	Lien, C., Liu, J., Liu, W., Louised, H., Lozard, R.J., Lu, X.,		
	Lucier, A., Lucier, R., Lunn, R., Ma, J., Maheshwari, M., Magua, P.,		
	Marandel, I., Martin, R., Martindale, A., Martinez, E., Massey, E.,		
	Mawney, E., McLeod, M.P., Meador, M., Mel, G., Metzger, S.,		
	Melzer, M., Miller, A., Miner, G., Miner, Z., Mitchell, T.,		
	Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M.,		
	Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,		
	Nguyen, N., Nickerson, E., Nwokwenkwo, S., Ogih, M., Okunou, G.,		
	Oragunye, N., Oyale, R., Pace, A., Payton, B., Peery, U., Perez, L.,		
	Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,		
	Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Saverly, G.,		
	Scherer, S., Scott, G., Shen, H., Shih, C., Shocher, I., Sison, I.,		
	Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,		
	Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tameris, K., Tang, H.,		
	Taney, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,		
	Usmani, K., Vaequez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,		
	Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,		
	Williams, G., Williamson, A., Wleciyk, R., Wooden, S., Worley, K.,		
	Mu, C., Mu, Y., Mu, Y.F., Zhou, J., Zorrilla, S., Kucherlapati, R.,		
	Weinstock, G. and Gibbs, R.		
	Direct Submission		
TITLE	Unpublished		
JOURNAL	2 (bases 1 to 125001)		
REFERENCE	Worley, K.C.		
AUTHORS	Direct Submission		
TITLE	Submitted (08-JUL-2002) Human Genome Sequencing Center, Department		
JOURNAL	of Molecular and Human Genetics, Baylor College of Medicine, One		
	Baylor Plaza, Houston, TX 77030, USA		
	3 (bases 1 to 125001)		
REFERENCE	Worley, K.C.		
AUTHORS	Direct Submission		
TITLE	Submitted (29-JUL-2002) Human Genome Sequencing Center, Department		
JOURNAL	of Molecular and Human Genetics, Baylor College of Medicine, One		
	Baylor Plaza, Houston, TX 77030, USA		
	4 (bases 1 to 125001)		
REFERENCE	Worley, K.C.		
AUTHORS	Direct Submission		
TITLE	Submitted (30-JUL-2002) Human Genome Sequencing Center, Department		
JOURNAL	of Molecular and Human Genetics, Baylor College of Medicine, One		
	Baylor Plaza, Houston, TX 77030, USA		
	On Jul 30, 2002 this sequence version replaced gi:22002091.		
COMMENT	INFORMATION: http://www.bgsc.bcm.tmc.edu/ or email gc-help@bcm.tmc.edu		

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green

Carron, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chiu, D., Chowdhry, I., Christopoulos, C., Cleveland, C. D., Cox, C., Coyle, M. D., Dabhome, S. R., David, R., Davila, M. L., Davis, C., Day-Carroll, L., Dedrich, D. A., Delaney, K. R., Delgado, O., Denn, A. L., Ding, Y., Dinh, H. H., Douthett, K. J., Draper, H., Dugan-Bochs, S., Durkin, K. J., Earhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Emwelling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Franz, P., Gabler, A., Gao, J. Y., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B., Homi, F., Howard, S., Huber, J., Huliy, S., Hume, J., Ioshikhes, I., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolyet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kravitz, J., Kureishi, A., Landry, N., Leal, B., Lee, E., Lewis, L. C., Lewis, L., Li, J., Li, Z., Lichard, O., Lien, C., Liu, J., Liu, W., Louiseged, H., Lozada, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapus, P., Macdonald, I., Martin, R., Martindale, A., Martinez, E., Massey, E., McWhiney, E., McLeod, M. P., Medor, M., Mei, G., Meschter, S., Metzker, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K. T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokemkwo, S., Ogih, M., Okumou, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y., Rivers, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shoohtari, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, K., Sutton, A. A., Swatek, A., Tabors, P., Tamezias, A., Tamezias, K., Tang, H., Tamei, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, Y., Villalon, D., Vinson, R., Wang, S., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wlezyrk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. P., Zhou, J., Zorrilla, S., Kuchertapatti, R., Weinstein, G., and Gibbs, R.

Unpublished

Direct Submission

2 (bases 1 to 125001)

Worley, K. C.

Direct Submission

Submitted (08-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 125001)

Worley, K. C.

Direct Submission

Submitted (23-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 125001)

Worley, K. C.

Direct Submission

Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

Bayor Plaza, Houston, TX 77030, USA

On Jul 30, 2002 this sequence version replaced gi:22002091.

INFORMATION: <http://www.hgsc.bcm.tmc.edu> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping listings.

ANNOTATION OF FEATURES:
STSs are identified using ePCR (Genome Reel, 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green

SOURCE Homo sapiens (human)
ORGANISM
REFERENCE Eukaryote, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
AUTHORS Mammalia, Eutheria, Primates, Catarrhini, Homnidae, Homo.
1 (bases 1 to 121520)
Hillier, L.W., Fulton, R.S., Fulton, L.A., Graves, T.A., Pepin, K.H.,
Wagner-McPherson, C., Layman, D., Maas, J., Jaeger, S., Walker, R.,
Wylie, K., Sekhon, M., Becker, M.C., O'Leahighlin, M.D., Schaller, M.E.,
Reuell, G.A., Delahunty, K.D., Miner, T.L., Nash, W.E., Cordes, M.,
Du, H., Sun, H., Edwards, J., Bradshaw-Cordum, H., Ali, J., Andrews, S.,
Isak, A., Vabnick, A., Nguyen, C., Du, F., Lamar, B., Courtney, L.,
Kalkbick, J., Ozerky, P., Biellecki, L., Scott, K., Holmes, A.,
Harkins, R., Harris, A., Strong, C.M., Hou, S., Tomlinson, C.,
Dauphin-Kohlberg, S., Kozlowicz-Reilly, A., Leonard, S., Rohlfing, T.,
Rock, S.M., Tin-Mollam, A.M., Abbott, A., Mink, P., Maupin, R.,
Stromwater, C., Latreille, P., Miller, N., Johnson, D., Murray, J.,
Moessner, J.P., Wendt, M.C., Yang, S.P., Schultz, B.R., Wallis, J.W.,
Sleeter, J., Bieri, T.A., Nelson, J.O., Berkowicz, N., Wohlmann, P.E.,
Cook, L.L., Hickenbotham, M.T., Eldred, J., Williams, D., Bedell, J.A.,
Mardis, E.R., Clifton, S.W., Chissee, S.L., Marr, M.A., Raymond, C.,
Haugen, E., Gillet, W., Zhou, Y., James, R., Phelps, K., Iadonoto, S.,
Bubb, K., Simms, E., Levy, R., Clendenning, J., Kaul, R., Kent, W.J.,
Porely, T.S., Baertsch, R.A., Brent, M.R., Kehler, E., Flick, P.,
Bork, P., Suyama, M., Bailey, J.A., Portnoy, M.E., Torrents, D.,
Chinwalla, A.T., Gibb, W.R., Eddy, S.R., McPherson, J.D., Olson, M.V.,
Elchler, E.E., Green, E.D., Waterston, R.H. and Wilson, R.K.
The DNA sequence of human chromosome 7
Nature 424 (6945), 157-164 (2003)
22737999
MEDLINE 12853948
PUBMED
REFERENCE 2 (bases 1 to 121520)
AUTHORS Bauer, C., Joshi, C. and Le, T.
TITLE The sequence of Homo sapiens PAC clone RP4-802G15
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 121520)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 121520)
Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 121520)
Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 121520)
Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (03-JUN-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
7 (bases 1 to 121520)
Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
8 (bases 1 to 121520)
Wilson, R.
TITLE Direct Submission
JOURNAL Submitted (27-JUN-2004) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Jul 3, 2001 this sequence version replaced gi:7243871.

Center: Washington University Genome Sequencing Center
Center code: MUGSC
Web site: http://genome.wustl.edu
Contact: saplens@wustl.wustl.edu

----- Summary Statistics
Center project name: H_DJ0802G15

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION: The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu>

SOURCE INFORMATION: This clone was derived from human PAC library RPC1-4, prepared by Pieter de Jong and coworkers at <http://www.chori.org> using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor. The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>) or from Pieter de Jong. VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the right is CTA-318M5, 200 bp overlap.
Actual start of this clone is at base position 1 of Rp4-802G315.
Actual end is at base position 14506 of CTA-318M5.

FEATURES	location/Qualifiers
source	1..121520
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="7"
	/map="7"
	/clone="RP4-802G15"
	/clone_1fb="RPC1-4"
	10..995
repeat_region	/rpt_family="L1"
	996..1816
repeat_region	/rpt_family="L1"
	1975..2278
repeat_region	/rpt_family="Alu"
	2821..2977
repeat_region	/rpt_family="L2"
	4195..4238
repeat_region	/rpt_family="L1"
	5073..5257
repeat_region	/rpt_family="MIR"
	5870..6457
repeat_region	/rpt_family="L1"
	6463..7323
repeat_region	/rpt_family="L1"
	7599..7638
repeat_region	/rpt_family="CT-rich"
	7667..7857
repeat_region	/rpt_family="MIR"
	8517..8547
repeat_region	/rpt_family="AT-rich"
	9124..9444

```

repeat_region      /rpt_family="L1"
                    9445..9796 "MALR"
repeat_region      /rpt_family="L1"
                    9797..10969
repeat_region      /rpt_family="L1"
                    10970..11273
repeat_region      /rpt_family="Alu"
                    11274..11395
repeat_region      /rpt_family="L1"
                    11414..12199
repeat_region      /rpt_family="L1"
                    18423..18683
repeat_region      /rpt_family="L1"
                    18676..18921
repeat_region      /rpt_family="L1"
                    18922..19856
repeat_region      /rpt_family="L1"
                    19983..20104
repeat_region      /rpt_family="L1"
                    20105..20470
repeat_region      /rpt_family="MALR"
                    20471..20591
repeat_region      /rpt_family="L1"
                    20592..20898
repeat_region      /rpt_family="Alu"
                    20902..20925
repeat_region      /rpt_family="(GAA)n"
                    20932..21065
repeat_region      /rpt_family="L1"
                    21079..22678
repeat_region      /rpt_family="L1"
                    22675..22901
repeat_region      /rpt_family="MALR"
                    22902..24732
repeat_region      /rpt_family="L1"
                    24733..24843
repeat_region      /rpt_family="Alu"
                    24844..25627
repeat_region      /rpt_family="L1"
                    25628..25734
repeat_region      /rpt_family="MER2_type"
                    25768..25823
repeat_region      /rpt_family="CT-rich"
                    25851..26075
repeat_region      /rpt_family="Alu"
                    26076..26684
repeat_region      /rpt_family="MER2_type"
                    26685..26719
repeat_region      /rpt_family="(CA)n"

Query Match      17.5%; Score 1507.4; DB 9; Length 121520;
Best Local Similarity 79.3%; Pred. No. 0;
Matches 2020; Conservative 0; Mismatches 361; Indels 166; Gaps 13;

3 TATGAAGCCATGTCTTTAATACCAAAACAGAAAGGATATA--CAAAAAGAAAAC 60
Db 92835 TATGAAGCTAGATCACTTAATACCAAAACAGAAAGGATACACCAAAAAGAAAAC 92894
Qy 61 TATGAAGCAGTACCACTGATGAATATATACATGCAAAATCCCAACAAATATAGCTAAC 120
Db 92895 TACAGGCTGATATCCCTGATGAACACAGAGCTAAATCTTAAACAAAATAGCTAGCTAAC 92954
Qy 121 CCAATCCACAGCATTTCAAGAAAGATATATCCACCATTTGCAATGGTTTCATACCAAGG 180
Db 92955 CCAATCCACCAATATCAAAAAGACAAATCCACCATGATCAAGTGGTTTCATACCAAGG 93014
Qy 181 GTGCA-GGATAGATTACATACCAAGTCAATTAATGTGTATACATACATTAACAGATT 239
Db 93015 ATGGAGGAGTGTATTACATATAGCAAGTCAATTAATGTGTATACATACATTAACAGATT 93074
Qy 240 AAAAAAATAATCAATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAATCTAA 299
Db 93075 AAAAAAATAATCAATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAATCTCAAG 93134

```

```

Qy 300 CATTTCTTATGATTTAAACCTTCAGCAAAATCGACATAGAAAGGACATACCTTAATGTA 359
Db 93135 CATCTTTATATATTTAAACCTTCAGCAAAATCGACATAGAAAGGACATACCTTAATGTA 93194
Qy 360 ATAAAAGCCATATATGACGGAACCAAGCAAACTTATATAGTAAGGGAAAAAGTTGAAA 419
Db 93195 ATAAAAGCCATCTATGACAAACCAAGTAACTATATAGTAAGGGAAAAAGTTGAAA 93254
Qy 420 ACATTGTCCTGAGAACTGGAACAAAGACAGATG-CTACTTCAACCACTTATTTCAAC 478
Db 93255 GCATTCCTCTGAGAAATGAAACAAAGACAGATGCCCTCTCTCAACCTCTTCAAC 93314
Qy 479 ATAGTAGTGAAGCTTTTACCCAGACAAATCAACAAGAAAGAAATGAAGGGACCCAA 538
Db 93315 ACAGTAGTGAAGCTCTTACCCAGACAAATCAACAAGAAAGAAATGAAGGGATTCGA 93374
Qy 539 ATCAATTAAGAGAAAGTCAAACTGTCCTGTTCACTGATGATGATGTTATACCTAGAA 598
Db 93375 ATTGATTAAGAGAAAGTCAATCTGTCACGTTTGTGATGATGATGTTATACCTCAA 93434
Qy 599 AACCTTAAGACTCATCCAGAAAGCTCCTGAACTGATACATTAATTCATAGTTTCA 658
Db 93435 AACCTTAAGACTCCTCTAGAAAGCTCCTGAACTGATTAAGATTCACAAAGTTTCC 93494
Qy 659 GGATCAAACTAAATGTACAAATCAAGTACGATGATATACAGCAAGTACCAAGCT 718
Db 93495 AGATCAAAATTAATGTACAAATCAAGTACGATGATGATGATGATGATGATGATGATG 93554
Qy 719 GGAATCAATCAAGAACTCAAACTTTTCAATAGCTGTAAGAAAAA-----TACTT 771
Db 93555 GGAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 93614
Qy 772 AAGATATTTCTTATCCCAAGAGGTGAAGGACCTCTTCAAGAAACTTCAAAACCACT 831
Db 93615 AAGATATTTCTTATCCCAAGAGGTGAAGGACCTCTTCAAGAAACTTCAAAACCACT 93674
Qy 832 ----GACATCATAGATGACAAACAAAGTGAACAACTCCATGCTCATGAGGGTAG 887
Db 93675 GAAAGAAATCATAGATGACAAACAAAGTGAACAACTCCATGCTCATGAGGGTAG 93734
Qy 888 AATCAATATTTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 947
Db 93735 AATCAATATTTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 93794
Qy 948 CACCAAAATATCATCTCTTCTTCAAGAACTTCAAAATTCATAT 1007
Db 93795 CATCAAAAACCAACATCATCTCTTCAAGAACTTCAAAATTCATAT 93853
Qy 1008 GGAACAAACCAAAAAAAGAAAAAAGCCGATAGCCAAAGCAAGCTTAGCAAAA 1067
Db 93854 GGAAC-----CAAAAAGAGTCCATAGCCAAAGCAAGCTTAGCAAAA 93898
Qy 1068 GAACAATCTGAGGATCATATCCATCTTCAACTATATCTAACAGGCTATATATCAAC 1127
Db 93899 GAACAATCTGAGGATCATATCTGATCTTCAACTATATCTTAAAGGCATGTCAC 93958
Qy 1128 CAAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1187
Db 93959 CAAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 94018
Qy 1188 TCCAGAAATTAAGCCAAATATATATAGCAACGATTTTGAAGAAAGCAAAACAT 1247
Db 94019 CCGAGAAATTAAGCCAAACATTAAGCAACGATTTTGAAGAAAGCAAAACAT 94078
Qy 1248 AAGTGGGAAAAGACATTTCTAGTTAACTAATGATGATGATGATGATGATGATGATGATG 1307
Db 94079 AAGTGGGAAAAGACATTTCTTAACTAATGATGATGATGATGATGATGATGATGATGATG 94138
Qy 1308 TGAAGATGAAGCTGATCCCTTGTCTTCACTTAATCAAAATTTGATACAAAGTGA 1367
Db 94139 TGAAGATGAAGCTGAT-CTCATCTCTCTCACTTATCAAAATTCACATCAAGATGA 94197

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: December 1, 2004, 23:06:03 ; Search time 3465 Seconds
(without alignments)
13062.196 Million cell updates/sec

Title: US-09-867-570-3

Perfect score: 8622

Sequence: 1 tgcctgaagccatgcact.....gtccccaagcccttacc 8622

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

N_Geneseq_23sep04:*
1: Geneseq19808:*
2: Geneseq19908:*
3: Geneseq20008:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8622	100.0	8622	6	ABK52823
2	1401	16.2	66494	6	ABO88140
3	1390.4	16.1	110000	6	ABR90193_0
4	1380.4	16.1	110000	6	ABQ87681_0
5	1390.4	16.1	110000	6	ABX33717_0
6	1375.6	16.0	110000	6	ABX08336_13
7	1375.6	16.0	110000	12	ADJ25985_13
8	1375.6	16.0	110000	12	ADN97989_13
9	1374.6	15.9	1400	6	ADJ33751
10	1374.6	15.9	1400	8	ABZ42595
11	1374.6	15.9	1400	12	ADH08534
12	1374	15.9	110000	12	ADH08534
13	1358	15.8	99014	6	ABN96931
14	1355	15.7	42738	4	AAK68992
15	1346	15.6	136284	6	ABK83575
16	1311.6	15.2	110000	4	AAK95240_05
17	1311.6	15.2	110000	4	AAK96733_05
18	1311.6	15.2	110000	6	ABT00010_05
19	1311.6	15.2	110000	6	ABT01503_05
20	1311.6	15.2	110000	12	ADH77485_05
21	1308	15.2	349881	10	ADC66642

C	22	1302	15.1	145606	12	ADP44617
	23	1301.8	15.1	32167	4	AAI05509
	24	1288.6	14.9	247682	12	ADL08109
	25	1286.8	14.9	148567	9	ACA62841
	26	1286.8	14.9	148567	10	ABR55500
	27	1286.8	14.9	148567	10	ABR55500
	28	1280.4	14.9	148567	6	ABT109163
	29	1280.4	14.9	70665	6	ABT10716
	30	1280.4	14.9	70665	11	ADN95672
	31	1274	14.8	1369	10	ADN95672
	32	1271.6	14.7	32249	4	AAI05336
	33	1271.6	14.7	32249	4	ABL98205
	34	1264	14.7	1604	6	ABD33752
	35	1264	14.7	1604	8	ABZ42602
	36	1247	14.5	79590	10	ADH08536
	37	1247	14.5	79590	11	ADL27152
	38	1247	14.5	79684	9	ADA03074
	39	1247	14.5	79684	9	ADA66358
	40	1245.6	14.4	136284	10	ADH72812
	41	1243.8	14.4	4019	10	AAK95601
	42	1228.4	14.2	2040	6	ADH08519
	43	1228.4	14.2	2040	12	ADH08519
	44	1228.4	14.2	2040	12	ADH08519
	45	1225.8	14.2	110000	12	ADH69807_0

ALIGNMENTS

RESULT 1	ABK52823	standard; DNA; 8622 BP.
ID	ABK52823	standard; DNA; 8622 BP.
XX	ABK52823	
AC	ABK52823	
XX	27-AUG-2002	(first entry)
DT	27-AUG-2002	(first entry)
XX	Genomic DNA encoding human G-protein coupled receptor (GPCR).	
DE	Human; G-protein coupled; receptor; GPCR; human protease;	
XX	human therapeutic protein; query sequence; search; gene; ds;	
KW	sequence database; non-human transgenic animal; gene therapy;	
KW	Chromosome 3.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX		
XX	Key	Location/Qualifiers
FT	CDS	4300..8496
FT		/*tag= a
FT		/product= "Human G-protein coupled receptor (GPCR)"
FT	exon	4300..4319
FT		/*tag= b
FT	intron	4320..7502
FT		/*tag= c
FT	exon	/number= 1
FT		7503..8496
FT		/*tag= d
FT		/number= 2
XX	W0200234914-A1.	
XX	02-MAY-2002.	
PD	10-OCT-2001; 2001WO-US031592.	
XX	25-OCT-2000; 2000US-00695045.	
PR	31-MAY-2001; 2001US-00867570.	
XX	(PEKE) PE CORP NY.	
PA	Wei M, Zhao Q, Woodage T, Di Francesco V, Beasley EM,	
XX		

DR WPI: 2002-463360/49.
DR P-PSDB: AAU97598.
PT Novel isolated G-protein coupled receptor peptide useful for treating
PT disorder characterized by absence of, in appropriate or unwanted
PT expression of the receptor protein, and as immunogens to raise
PT antibodies.
XX
PS Claim 4; Fig 3; 75bp; English.
XX
CC The present invention relates to a new G-protein coupled receptor (GPCR)
CC peptide. The invention is useful for identifying a modulator of GPCR and
CC for treating a disease or condition mediated by a human protease. The
CC invention is also useful as models for the development of human
CC therapeutics, for identifying therapeutic proteins, as targets for
CC development of human therapeutic agents, and as query sequence to perform
CC a search against sequence databases to, for e.g., identify other family
CC members of related sequences. The vector of the invention is useful for
CC producing a GPCR protein or peptide, for conducting cell-based assays
CC involving the GPCR protein or its fragment, for identifying GPCR protein
CC mutants whose functions are affected, and to produce non-human transgenic
CC animals. The present nucleic acid sequence represents the human G-protein
CC coupled receptor (GPCR) gene located on chromosome 3. This sequence
CC encodes the human G-protein coupled receptor (GPCR) protein of the
CC invention
XX
SQ Sequence 8622 BP; 2684 A; 1921 C; 1679 G; 2338 T; 0 U; 0 Other;
Query Match 100.0%; Score 8622; DB 6; Length 8622;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 8622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTATGAAGCCCATGCTCTTATATCCAAAACGAGAAAGATATTCAAAAGAAAC 60
DB 1 TGTATGAAGCCCATGCTCTTATATCCAAAACGAGAAAGATATTCAAAAGAAAC 60
QY 61 TATAGACCGATGACCACTGATGATATATACATGAGAAATCCCAACAAATAGCTAAC 120
DB 61 TATAGACCGATGACCACTGATGATATATACATGAGAAATCCCAACAAATAGCTAAC 120
QY 121 CCAATCCACAGCATATCAAGAAAGATATTCACCATTTGCAAGTGGGTTTCATACCAAGG 180
DB 121 CCAATCCACAGCATATCAAGAAAGATATTCACCATTTGCAAGTGGGTTTCATACCAAGG 180
QY 181 GTGCGAGATAGTTTAACTATACCAAGTCAATTAATGATGATCACTAAAGAAATTA 240
DB 181 GTGCGAGATAGTTTAACTATACCAAGTCAATTAATGATGATCACTAAAGAAATTA 240
QY 241 AAAACAAAATACATGATCATCTCAATAGATCTGAAAAAGATTGACAAATCTAAC 300
DB 241 AAAACAAAATACATGATCATCTCAATAGATCTGAAAAAGATTGACAAATCTAAC 300
QY 301 ATTTCTTTATGATTAACCTTCAGCAAAATGACATAGAAAGACATACCTTAATGTAA 360
DB 301 ATTTCTTTATGATTAACCTTCAGCAAAATGACATAGAAAGACATACCTTAATGTAA 360
QY 361 TAAAAGCCATATATGACGAGCCACAGCAAACTTTATCTGAATGGGAAAAATTTAAAA 420
DB 361 TAAAAGCCATATATGACGAGCCACAGCAAACTTTATCTGAATGGGAAAAATTTAAAA 420
QY 421 CATTTGCTCTGAGAACTGGAACAAGAGATGCTACTTCAACCACTTCTATTCAACAT 480
DB 421 CATTTGCTCTGAGAACTGGAACAAGAGATGCTACTTCAACCACTTCTATTCAACAT 480
QY 481 AGTAGTGAAGTTTATGCGAGACATCAAGCAAGAAAGAAATCAAGGGCAACCAAT 540
DB 481 AGTAGTGAAGTTTATGCGAGACATCAAGCAAGAAAGAAATCAAGGGCAACCAAT 540
QY 541 CAATTAAGGAGAGTAACTGCTCTGCTTCACTGATGATGATGATGATGATGATGATGAT 600
DB 541 CAATTAAGGAGAGTAACTGCTCTGCTTCACTGATGATGATGATGATGATGATGATGAT 600
QY 601 CCTTAAGAGTCAATCCAGAAAGCTCTAGAACTGATACATTAATTCAGTAAAGTTTCAGG 660

DB 601 CCTTAAGAGTCAATCCAGAAAGCTCTAGAACTGATACATTAATTCAGTAAAGTTTCAGG 660
QY 661 ATACAAACTAAATGTACACAAATCAATGACCTGCTATACCAACAGTGACCAAGCTGA 720
DB 661 ATACAAACTAAATGTACACAAATCAATGACCTGCTATACCAACAGTGACCAAGCTGA 720
QY 721 GAATCAAACTCAAGAACTCAAACTTTTCAATCTGTAATAAATCTTAAGAAATTT 780
DB 721 GAATCAAACTCAAGAACTCAAACTTTTCAATCTGTAATAAATCTTAAGAAATTT 780
QY 781 CTACCCAGAGAGTGAAGAGCCTCTACAGAGAAATCTACAAACAGCTGACATCTATA 840
DB 781 CTACCCAGAGAGTGAAGAGCCTCTACAGAGAAATCTACAAACAGCTGACATCTATA 840
QY 841 GATGACACAAACAGTGGAAACAATCCATGCTCATGATGGTGAATCAATATTTGTG 900
DB 841 GATGACACAAACAGTGGAAACAATCCATGCTCATGATGGTGAATCAATATTTGTG 900
QY 901 AAAATGACATATTTGCAAGCAATCTTCAAGTTCATGCAATTCACCAACCAATATCA 960
DB 901 AAAATGACATATTTGCAAGCAATCTTCAAGTTCATGCAATTCACCAACCAATATCA 960
QY 961 TCATCATTTCTGACAGAACTAGAAAAAAACAATTTCAATATGGAACCAACCA 1020
DB 961 TCATCATTTCTGACAGAACTAGAAAAAAACAATTTCAATATGGAACCAACCA 1020
QY 1021 AAAAAAAAAAAAAAAAAAATCCGCTAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1080
DB 1021 AAAAAAAAAAAAAAAAAAATCCGCTAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1080
QY 1081 GGCATCACTTCCCATCTTCAACATATCTCAAGGCTATATGACCAAAACATCATAG 1140
DB 1081 GGCATCACTTCCCATCTTCAACATATCTCAAGGCTATATGACCAAAACATCATAG 1140
QY 1141 CACTGACATTAACCTAGGACATAGCAATGGAAGAAAGAAAGAAAGAAAGAAAGAA 1200
DB 1141 CACTGACATTAACCTAGGACATAGCAATGGAAGAAAGAAAGAAAGAAAGAAAGAA 1200
QY 1201 CCAATTAATTAATGCACTGATTTTGAACAAGCAACAAACATTAAGTGGGAAAA 1260
DB 1201 CCAATTAATTAATGCACTGATTTTGAACAAGCAACAAACATTAAGTGGGAAAA 1260
QY 1261 GACATTTCTATGATTAACAAATGAGTGAATTTTGAACAAGCAACATGAGGAAGTAAA 1320
DB 1261 GACATTTCTATGATTAACAAATGAGTGAATTTTGAACAAGCAACATGAGGAAGTAAA 1320
QY 1321 CTGATCCCTGTCTCTCACTTAATCAAAAAATGATCAAGATGATCAAGACTTAA 1380
DB 1321 CTGATCCCTGTCTCTCACTTAATCAAAAAATGATCAAGATGATCAAGACTTAA 1380
QY 1381 TCTGAGACTTAAACCATTAATAATCTGAGATTAACATCAGAAAAATGCTTGAACAT 1440
DB 1381 TCTGAGACTTAAACCATTAATAATCTGAGATTAACATCAGAAAAATGCTTGAACAT 1440
QY 1441 TCACCTAGGCAAGAACTTATGAGCCAAAGAACCAAAAGTAAATGCAACAAACAAAT 1500
DB 1441 TCACCTAGGCAAGAACTTATGAGCCAAAGAACCAAAAGTAAATGCAACAAACAAAT 1500
QY 1501 AATATGATGAGCTTAATTAATTAACCTAAAGCTTTTGGCAGCAAAACATCATAGCAG 1560
DB 1501 AATATGATGAGCTTAATTAATTAACCTAAAGCTTTTGGCAGCAAAACATCATAGCAG 1560
QY 1561 AGCAACAGACAAACCAAGTGAAGAAATCTTCAACAACTAGACATCTGACTAAG 1620
DB 1561 AGCAACAGACAAACCAAGTGAAGAAATCTTCAACAACTAGACATCTGACTAAG 1620
QY 1621 ACTAATATCCGGAATTCACAAGGAACTCAACAAATCGCAAGAGAAAGCAACATCC 1680
DB 1621 ACTAATATCCGGAATTCACAAGGAACTCAACAAATCGCAAGAGAAAGCAACATCC 1680
QY 1681 CATGAAAGAGTGGGCTTAAGACATGAATGACAAATTCCTCAAAAGAAAGATATCAATGSC 1740

```
Db 1681 CATGAAGATGGGCTAAGACATGATAGACAAATTCCTCAAAAGAGATATACAAATGGC 1740
Qy 1741 CAACAAACAGGAAAAATGCTTAAACATCATATGATTTAGGAAATGTAATCAACACTG 1800
Db 1741 CAACAAACAGGAAAAATGCTTAAACATCATATGATTTAGGAAATGTAATCAACACTG 1800
Qy 1801 TTAATGCGATACACCTTATCTCTGCAAGATGTCATATTTTAAATCTTAAATTAATA 1860
Db 1801 TTAATGCGATACACCTTATCTCTGCAAGATGTCATATTTTAAATCTTAAATTAATA 1860
Qy 1861 GATGTGGTGGGTCTGTGTGATTAAGAGACATTTTACACTGCTGTGGGATGTAAAC 1920
Db 1861 GATGTGGTGGGTCTGTGTGATTAAGAGACATTTTACACTGCTGTGGGATGTAAAC 1920
Qy 1921 TTGGCGAACCATCTATGAAAAAGGTGGAATTTCTTAAGAACTTAAAGTGAATGAC 1980
Db 1921 TTGGCGAACCATCTATGAAAAAGGTGGAATTTCTTAAGAACTTAAAGTGAATGAC 1980
Qy 1981 CATTTGATCCAGCAATCCATTAATATGTAATAATATATATTTATATATCCATGAAAT 2040
Db 1981 CATTTGATCCAGCAATCCATTAATATGTAATAATATATATTTATATATCCATGAAAT 2040
Qy 2041 ACAACTCAGCCATAAAAAGATTAATGATGACATTCACAGCAATCTAGATGGAATGG 2100
Db 2041 ACAACTCAGCCATAAAAAGATTAATGATGACATTCACAGCAATCTAGATGGAATGG 2100
Qy 2101 AGACCTTTATTTCTAAGTGGGGTAACTCAGAAATGAAAAACCAACATCATATGTTCTAC 2160
Db 2101 AGACCTTTATTTCTAAGTGGGGTAACTCAGAAATGAAAAACCAACATCATATGTTCTAC 2160
Qy 2161 TTAAAGTGGGGGCTAAGCTGTGAGGACACAGAGGCAATGAAATGATATATATGAACTG 2220
Db 2161 TTAAAGTGGGGGCTAAGCTGTGAGGACACAGAGGCAATGAAATGATATATATGAACTG 2220
Qy 2221 GGACTTGAAGGGAAGATGGAAGAGAGGCGAGGATAAAGACTACACATGGGTACAGT 2280
Db 2221 GGACTTGAAGGGAAGATGGAAGAGAGGCGAGGATAAAGACTACACATGGGTACAGT 2280
Qy 2281 GTACACTGCTCAGGTGATGGGTGACCAAAATCTCAGAAATTCACATAAGAACTTATC 2340
Db 2281 GTACACTGCTCAGGTGATGGGTGACCAAAATCTCAGAAATTCACATAAGAACTTATC 2340
Qy 2341 CATGGAAGCAACACCCCTGTTCCCAAAATCCCAAGAAATTAATAATTAATAATA 2400
Db 2341 CATGGAAGCAACACCCCTGTTCCCAAAATCCCAAGAAATTAATAATTAATAATA 2400
Qy 2401 ATGATTTAATTTCAAGAAATTTAAAAAGTCACTGTTCAAGTTTATATATATGAAGTA 2460
Db 2401 ATGATTTAATTTCAAGAAATTTAAAAAGTCACTGTTCAAGTTTATATATATGAAGTA 2460
Qy 2461 AGAATGAAAAAGTGTAGCAAGTGTAGCCTCTGCACAATGGGACTCTAGATTTTCACTTG 2520
Db 2461 AGAATGAAAAAGTGTAGCAAGTGTAGCCTCTGCACAATGGGACTCTAGATTTTCACTTG 2520
Qy 2521 CATACACTTCTCTGCAATTTGAAAGAAAGTATACATGAAATATATACACATATGATA 2580
Db 2521 CATACACTTCTCTGCAATTTGAAAGAAAGTATACATGAAATATATACACATATGATA 2580
Qy 2581 AAGAAAAACATCAAAAAATGTGTAGAGGCAATGTAGCCTGAAATGGTCCCATGATCTAC 2640
Db 2581 AAGAAAAACATCAAAAAATGTGTAGAGGCAATGTAGCCTGAAATGGTCCCATGATCTAC 2640
Qy 2641 TTTTTCATTTGATATATAAGCCTCATATGATAGTTCATTTGCTTAATGTATGCCATG 2700
Db 2641 TTTTTCATTTGATATATAAGCCTCATATGATAGTTCATTTGCTTAATGTATGCCATG 2700
Qy 2701 GCCCATATTTGATTTTAAATCAGACAGCAATTAATTAACAGAAAGTTGAACAAGATGG 2760
Db 2701 GCCCATATTTGATTTTAAATCAGACAGCAATTAATTAACAGAAAGTTGAACAAGATGG 2760
Qy 2761 GACGTGATAGAGAGGCTTAATGATATGATATGAGACAGAGGCAAGATCATCTCAG 2820
Db 2761 GACGTGATAGAGAGGCTTAATGATATGATATGAGACAGAGGCAAGATCATCTCAG 2820
Qy 2821 TTAGATTTGTGTCTCAAAATCTCTGCGCTGTGATTTGCCATAGTCTCATACAGAA 2880
Db 2821 TTAGATTTGTGTCTCAAAATCTCTGCGCTGTGATTTGCCATAGTCTCATACAGAA 2880
Qy 2881 ATTAACAGACTGTACACATCTTGTAAAGCTGTGATTTGCTACAGAGCTTTCATTTCACT 2940
Db 2881 ATTAACAGACTGTACACATCTTGTAAAGCTGTGATTTGCTACAGAGCTTTCATTTCACT 2940
Qy 2941 CCTGTAGGCACTCTCTGAATTTAGCAACAGAAAGTCTCTGAAGTCACTGAATCCA 3000
Db 2941 CCTGTAGGCACTCTCTGAATTTAGCAACAGAAAGTCTCTGAAGTCACTGAATCCA 3000
Qy 3001 GAAAGGCTCTCTACCTTTAGCAACAGGAGGTCTTCACTGAGCAAAAGAAAGAT 3060
Db 3001 GAAAGGCTCTCTACCTTTAGCAACAGGAGGTCTTCACTGAGCAAAAGAAAGAT 3060
Qy 3061 AAGGATAGTACCAAGAACTCTCTTCCACAGTCAAGTATGATTTTGTGTGAATC 3120
Db 3061 AAGGATAGTACCAAGAACTCTCTTCCACAGTCAAGTATGATTTTGTGTGAATC 3120
Qy 3121 ATGTCTTATGCTTCCACTTGTGTCTACATGACAGGGGTCAAGGCTTGTTCAGGAAA 3180
Db 3121 ATGTCTTATGCTTCCACTTGTGTCTACATGACAGGGGTCAAGGCTTGTTCAGGAAA 3180
Qy 3181 AGACAGAGACATGAAAGCTTCTTTCAGAAACTGATGCTGTCAACCCAACTGTGTAG 3240
Db 3181 AGACAGAGACATGAAAGCTTCTTTCAGAAACTGATGCTGTCAACCCAACTGTGTAG 3240
Qy 3241 CTCCTAATGTGTCCTCCCTCTTAAATTAATCTCCCAATACACTCTCTCATTTCCATCA 3300
Db 3241 CTCCTAATGTGTCCTCCCTCTTAAATTAATCTCCCAATACACTCTCTCATTTCCATCA 3300
Qy 3301 TTCAATCTGCTCATGAGAGAGCTGCTGCTCTTACATTCATTTAAGAGCAAGGAGAC 3360
Db 3301 TTCAATCTGCTCATGAGAGAGCTGCTGCTCTTACATTCATTTAAGAGCAAGGAGAC 3360
Qy 3361 ATGCAAGCAATTTCTTCCAGAGTTGAACCTGTATAGAGCAAGTTCTTGTGTTCACTTAC 3420
Db 3361 ATGCAAGCAATTTCTTCCAGAGTTGAACCTGTATAGAGCAAGTTCTTGTGTTCACTTAC 3420
Qy 3421 TTTTCAAAATTAATCTCTTCTTGTGCTATCTGAAAGGTCTAAGAAATATAGAGGCCA 3480
Db 3421 TTTTCAAAATTAATCTCTTCTTGTGCTATCTGAAAGGTCTAAGAAATATAGAGGCCA 3480
Qy 3481 ATTAATTAAGAGTGTTCATGAGAAAGTATTTAACAAGATGACAGAGTTAAGGCTCAG 3540
Db 3481 ATTAATTAAGAGTGTTCATGAGAAAGTATTTAACAAGATGACAGAGTTAAGGCTCAG 3540
Qy 3541 GATCTTAAGCAGCAATACATAGGGGAGCACTACTTCTCCCTTGAAGCGAGCAGG 3600
Db 3541 GATCTTAAGCAGCAATACATAGGGGAGCACTACTTCTCCCTTGAAGCGAGCAGG 3600
Qy 3601 GAAGAGACATTAACATTTGTGTGCAATAGCCATAGCTGTAGCCATTAAGGTTGGAGACAT 3660
Db 3601 GAAGAGACATTAACATTTGTGTGCAATAGCCATAGCTGTAGCCATTAAGGTTGGAGACAT 3660
Qy 3661 GAGCAGGCAAGTGAAGAAAGCTGCTGTGCAAGCAGCAGCAGCAGCAGGCTGATATAGTT 3720
Db 3661 GAGCAGGCAAGTGAAGAAAGCTGCTGTGCAAGCAGCAGCAGCAGCAGGCTGATATAGTT 3720
Qy 3721 TGATCTGTGTCTCCACAAATCTCATGTGATGTATTTCCAAATGTTGGAAGAAAGG 3780
Db 3721 TGATCTGTGTCTCCACAAATCTCATGTGATGTATTTCCAAATGTTGGAAGAAAGG 3780
Qy 3781 CCTGTGAGAGATATTAATGATCAAGGGAGTGTTCATGAAATGTTTAAACACAT 3840
Db 3781 CCTGTGAGAGATATTAATGATCAAGGGAGTGTTCATGAAATGTTTAAACACAT 3840
Qy 3841 CCCCTTTGATATGTTGTGTGATACAGAGTCTCATGAAATCTAGTGTGTTTAAA 3900
Db 3841 CCCCTTTGATATGTTGTGTGATACAGAGTCTCATGAAATCTAGTGTGTTTAAA 3900
```

QY 3901 GCGTGTAGACCTCCCTCTCTCTTACTCTGCTCTACAGATGTGAGAGCGCTGCTC 3960
 Db 3901 GCGTGTAGACCTCCCTCTCTCTTACTCTGCTCTACAGATGTGAGAGCGCTGCTC 3960
 QY 3961 CCCCTTTGCTTTTACAGAGATTTGGAAGCTTCTGAGGCTCTCCAGAGAGAGAGCTGC 4020
 Db 3961 CCCCTTTGCTTTTACAGAGATTTGGAAGCTTCTGAGGCTCTCCAGAGAGAGAGCTGC 4020
 QY 4021 TATGCTCTCTGTATAGCTGTAGAGCTATTAGCCAGTAAACCATTTCTCTATAAAT 4080
 Db 4021 TATGCTCTCTGTATAGCTGTAGAGCTATTAGCCAGTAAACCATTTCTCTATAAAT 4080
 QY 4081 TCCCACTCTCAGGATTTCTTTTATGCAATTGAGATGAACTAATATACAGACAGAGAG 4140
 Db 4081 TCCCACTCTCAGGATTTCTTTTATGCAATTGAGATGAACTAATATACAGACAGAGAG 4140
 QY 4141 CCAGAGATGAAATCCCAAGGTCTTCTGCTGTCTTCCAGTCTCTGCTGCTGCTC 4200
 Db 4141 CCAGAGATGAAATCCCAAGGTCTTCTGCTGTCTTCCAGTCTCTGCTGCTGCTC 4200
 QY 4201 CCAGGTCTCAATTTCCACAGAAACAGAAATTAAGAAATCCCACTGATGTGTACATA 4260
 Db 4201 CCAGGTCTCAATTTCCACAGAAACAGAAATTAAGAAATCCCACTGATGTGTACATA 4260
 QY 4261 GAAGCACTCTCTGAGATGTCAACAGAGTAAAGAAAGATGAGAAACAAATCTCATG 4320
 Db 4261 GAAGCACTCTCTGAGATGTCAACAGAGTAAAGAAAGATGAGAAACAAATCTCATG 4320
 QY 4321 TAAATGAGATCTCTCTCTCACTTCTTGTATCTCTCTAAATCTCTGAGGCTTCTCTATC 4380
 Db 4321 TAAATGAGATCTCTCTCTCACTTCTTGTATCTCTCTAAATCTCTGAGGCTTCTCTATC 4380
 QY 4381 TGATGATCCGCTCTCATTTGAGCTCTATCAGACTAATTAATGTTGGCTGTCTGCTC 4440
 Db 4381 TGATGATCCGCTCTCATTTGAGCTCTATCAGACTAATTAATGTTGGCTGTCTGCTC 4440
 QY 4441 TCTACTGCTCTTTATGAGAAATGTTGCTTTGTTAAATGAGTAAATTAATTAAT 4500
 Db 4441 TCTACTGCTCTTTATGAGAAATGTTGCTTTGTTAAATGAGTAAATTAATTAAT 4500
 QY 4501 GTTAATTTTAAAGAAATATGATGATTTGTTTGAATATTAAGTTGGCTGATCTAATAA 4560
 Db 4501 GTTAATTTTAAAGAAATATGATGATTTGTTTGAATATTAAGTTGGCTGATCTAATAA 4560
 QY 4561 GACATGAGAGAAATATCTTAAACAAGAAATGTTGGCTGCTGAGGCTGCTAGGT 4620
 Db 4561 GACATGAGAGAAATATCTTAAACAAGAAATGTTGGCTGCTGAGGCTGCTAGGT 4620
 QY 4621 CTGAATCTACAGATTTCAACAACTACAGAGAACTTTTCAAAAAATTAAGGTGTGCG 4680
 Db 4621 CTGAATCTACAGATTTCAACAACTACAGAGAACTTTTCAAAAAATTAAGGTGTGCG 4680
 QY 4681 GAGTGTGTATGTACGAAACAGGTACAACTTGTATTTCTTGTCAATTAATTTGAAAA 4740
 Db 4681 GAGTGTGTATGTACGAAACAGGTACAACTTGTATTTCTTGTCAATTAATTTGAAAA 4740
 QY 4741 CTACATATACAGAACTTATATAGATTTGATTTGTCAATTAATTTGAAAACTTT 4800
 Db 4741 CTACATATACAGAACTTATATAGATTTGATTTGTCAATTAATTTGAAAACTTT 4800
 QY 4801 AAATGATTTATGTATCTGGAGAAAGTGTAGATGATATACAAATCCATATATAGG 4860
 Db 4801 AAATGATTTATGTATCTGGAGAAAGTGTAGATGATATACAAATCCATATATAGG 4860
 QY 4861 AAATGAGATCTGAGATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4920
 Db 4861 AAATGAGATCTGAGATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4920
 QY 4921 AACAAAAATGACACTCTTGAAGATCTGAACTGAAGATCCAAAGATCTATACATGAG 4980
 Db 4921 AACAAAAATGACACTCTTGAAGATCTGAACTGAAGATCCAAAGATCTATACATGAG 4980
 QY 4981 TTCAAAAAATGCTGCTCCCAAGTCTTAGAGATGCTGCTCATCTTGTGTATCTAGAT 5040

Db 4981 TTCAAAAAATGCTGCTCCCAAGTCTTAGAGATGCTGCTCATCTTGTGTATCTAGAT 5040
 QY 5041 GGTTCACAGGACATTAAGATCTCACTTATGAAAAAAGACAGAGGAGAGAGGCT 5100
 Db 5041 GGTTCACAGGACATTAAGATCTCACTTATGAAAAAAGACAGAGGAGAGAGGCT 5100
 QY 5101 TTGCTCTCTTATTAATCCCATGAGCAGAGATCTGCTCTGCACTTTTGTGATCTTCC 5160
 Db 5101 TTGCTCTCTTATTAATCCCATGAGCAGAGATCTGCTCTGCACTTTTGTGATCTTCC 5160
 QY 5161 ACTTAAACAGACCTGCTCAATGAGATGTCAATCCAGATCAAGAAAACTGGATGTGCTC 5220
 Db 5161 ACTTAAACAGACCTGCTCAATGAGATGTCAATCCAGATCAAGAAAACTGGATGTGCTC 5220
 QY 5221 CTGTGTCTCTGTATCAATCTCAGAAAGGTTATGTGACCAAAAAAGAAATCTTGGGCA 5280
 Db 5221 CTGTGTCTCTGTATCAATCTCAGAAAGGTTATGTGACCAAAAAAGAAATCTTGGGCA 5280
 QY 5281 ACCAGAGCTCTCTGAGCCCTGAGCTGTCTGATTTCTGTGCTCATCAAGATTTTCA 5340
 Db 5281 ACCAGAGCTCTCTGAGCCCTGAGCTGTCTGATTTCTGTGCTCATCAAGATTTTCA 5340
 QY 5341 GGAATCTCTCAGAAATTAATTAATGTTGGGAGAGAACTGAGATCTTGTGACAGA 5400
 Db 5341 GGAATCTCTCAGAAATTAATTAATGTTGGGAGAGAACTGAGATCTTGTGACAGA 5400
 QY 5401 CTCAGAGGACAGAGGCTGTATGTGACCTGCTCTCAATGTTGTGAAACAGAAAACT 5460
 Db 5401 CTCAGAGGACAGAGGCTGTATGTGACCTGCTCTCAATGTTGTGAAACAGAAAACT 5460
 QY 5461 TTAATCTCTAGGCTTATGCTCATCTTAATGTTATATAGATTAATACATAGACGTCT 5520
 Db 5461 TTAATCTCTAGGCTTATGCTCATCTTAATGTTATATAGATTAATACATAGACGTCT 5520
 QY 5521 TTAAGAACATCATAGCATGTTAAACACATCTAAATGTTGTGATACACAGTGA 5580
 Db 5521 TTAAGAACATCATAGCATGTTAAACACATCTAAATGTTGTGATACACAGTGA 5580
 QY 5581 AGACAGGACATGACTTACTCTCTTAAGGATCTTGGGTTTCAATGAGAAACAACTATCA 5640
 Db 5581 AGACAGGACATGACTTACTCTCTTAAGGATCTTGGGTTTCAATGAGAAACAACTATCA 5640
 QY 5641 TACCATACCTATAGATGAGCAACAGTTAGTGTCTGAGTGTGATTAACAGAGTTCTC 5700
 Db 5641 TACCATACCTATAGATGAGCAACAGTTAGTGTCTGAGTGTGATTAACAGAGTTCTC 5700
 QY 5701 CTTTCTCTCCATTTCTTTTGGGCAATCAGAGCTGTGTGCTCTTAAGAG 5760
 Db 5701 CTTTCTCTCCATTTCTTTTGGGCAATCAGAGCTGTGTGCTCTTAAGAG 5760
 QY 5761 AGCTATGATGATGATGATCACTCTGATGCTCTCTTAATCTCCAGAGAGAGATGATC 5820
 Db 5761 AGCTATGATGATGATGATCACTCTGATGCTCTCTTAATCTCCAGAGAGAGATGATC 5820
 QY 5821 TTCTTCCACCTGAGAGGCTCTGAGATGCTCTTGAATCTCCAGAGAGAGATGATC 5880
 Db 5821 TTCTTCCACCTGAGAGGCTCTGAGATGCTCTTGAATCTCCAGAGAGAGATGATC 5880
 QY 5881 CCTGTATGAGCAAAAAAGAACTCTGAATTTGTTCTTAATGAGACGACTCACCTCTA 5940
 Db 5881 CCTGTATGAGCAAAAAAGAACTCTGAATTTGTTCTTAATGAGACGACTCACCTCTA 5940
 QY 5941 TTTTTCCTTATTTCAATTTGCTTCTCAATCTCTATCTGAGATTTTAAAGTTATTTT 6000
 Db 5941 TTTTTCCTTATTTCAATTTGCTTCTCAATCTCTATCTGAGATTTTAAAGTTATTTT 6000
 QY 6001 TTTTTCAGCCCAATTTTGAATGATCTGATCAATTTGAATTTGAATCACTCTCTACT 6060
 Db 6001 TTTTTCAGCCCAATTTTGAATGATCTGATCAATTTGAATTTGAATCACTCTCTACT 6060
 QY 6061 TTAACCCCTCTAATGATGATTAATGAGACATAGTGTGTGTGCTCAAGGCTGTGTA 6120

Db 6061 TTACCCCTCTAATCATGTATATATGACACATAGTGTGCTGGCTCCAAAGGCTGTGA 6120
Qy 6121 AAAAATGATCATGATGACGCTGTGCTGGCTCACTTCAATATATATATATATACCC 6180
Db 6121 AAAAATGATCATGATGACGCTGTGCTGGCTCACTTCAATATATATATATATACCC 6180
Qy 6181 CCGGGAAGAAATATCTCTCTTATATCTAATATTTCAATATCTGCAATACAGAAATTA 6240
Db 6181 CCGGGAAGAAATATCTCTCTTATATCTAATATTTCAATATCTGCAATACAGAAATTA 6240
Qy 6241 TACTAATTTTACCTCTAGGCTCTTCAATGATTAAGAGGCAATACCTAATTAATCTG 6300
Db 6241 TACTAATTTTACCTCTAGGCTCTTCAATGATTAAGAGGCAATACCTAATTAATCTG 6300
Qy 6301 TCAATCAGCTGTGTATTTCTCCAAATTAGACCTAATCTCAATCTTCCAGTTGAATTT 6360
Db 6301 TCAATCAGCTGTGTATTTCTCCAAATTAGACCTAATCTCAATCTTCCAGTTGAATTT 6360
Qy 6361 GCATGATATCTCTCTTACAAACCAAGCCCTACACTTCTCTATTTCCACTCATGACT 6420
Db 6361 GCATGATATCTCTCTTACAAACCAAGCCCTACACTTCTCTATTTCCACTCATGACT 6420
Qy 6421 CCTCTCATCAATATGTTTGCATCAACAAAGAAAGCTACAAAGATCTCCGAAAGAG 6480
Db 6421 CCTCTCATCAATATGTTTGCATCAACAAAGAAAGCTACAAAGATCTCCGAAAGAG 6480
Qy 6481 AATGAAATAGGTTTACATTTGTATCTCAGCAGAAACCTTAGTAGTCCCAATACATAT 6540
Db 6481 AATGAAATAGGTTTACATTTGTATCTCAGCAGAAACCTTAGTAGTCCCAATACATAT 6540
Qy 6541 TCCCACTTCAATTAATCTGTGCTGAGTGGCACTCAGGCTCACCTCACTTCTTCTTCT 6600
Db 6541 TCCCACTTCAATTAATCTGTGCTGAGTGGCACTCAGGCTCACCTCACTTCTTCTTCT 6600
Qy 6601 CTGTTCTATGCTGAGCAATCAGCTCAGACCAACCTTACCCAAACATCTGTGAACAA 6660
Db 6601 CTGTTCTATGCTGAGCAATCAGCTCAGACCAACCTTACCCAAACATCTGTGTAA 6660
Qy 6661 ATGCTTCTAGGGGTTTGGGCAAGCACTGAGTCTTATTTTAAAGCAATCAGTGTGT 6720
Db 6661 ATGCTTCTAGGGGTTTGGGCAAGCACTGAGTCTTATTTTAAAGCAATCAGTGTGT 6720
Qy 6721 CAATTTGAGTTTGGGCACTCATCAATCATTTCTTCAACACAGATGAGCTGTCCA 6780
Db 6721 CAATTTGAGTTTGGGCACTCATCAATCATTTCTTCAACACAGATGAGCTGTCCA 6780
Qy 6781 AATGAAATCTGATGAAATGAAATTTTCTCATCTAATATATGATGTGTCTAATGCT 6840
Db 6781 AATGAAATCTGATGAAATGAAATTTTCTCATCTAATATATGATGTGTCTAATGCT 6840
Qy 6841 TACATTTGCTTTCAATTTTATTTTCAATTTCAATCAATCTACATTTGCCATTAGGCTT 6900
Db 6841 TACATTTGCTTTCAATTTTATTTTCAATTTCAATCAATCTACATTTGCCATTAGGCTT 6900
Qy 6901 CTGATGATGATCTTCAATTTGATGAAAGCTTTATGAAAGCACTTTGTCTGTATG 6960
Db 6901 CTGATGATGATCTTCAATTTGATGAAAGCTTTATGAAAGCACTTTGTCTGTATG 6960
Qy 6961 GAATAGGCACTAGAGATATATAATGTAAATGTGCTGTCTGTCAATGATGACACT 7020
Db 6961 GAATAGGCACTAGAGATATATAATGTAAATGTGCTGTCTGTCAATGATGACACT 7020
Qy 7021 GAGTTATTTCTGACCCACAGGCTCCGCAATTTTCAACATCTTAGGCAAGATCCATT 7080
Db 7021 GAGTTATTTCTGACCCACAGGCTCCGCAATTTTCAACATCTTAGGCAAGATCCATT 7080
Qy 7081 TCCCTGCTGATCAATATGATGATCTTTTCTGTCTGAGAGATGACAGTCTGTGATG 7140
Db 7081 TCCCTGCTGATCAATATGATGATCTTTTCTGTCTGAGAGATGACAGTCTGTGATG 7140
Qy 7141 AGGCTGTCAACAACCTCTTGTGTATCTGAATTTCTCCACCTGAGAGAAATTTTCAG 7200
Db 7141 AGGCTGTCAACAACCTCTTGTGTATCTGAATTTCTCCACCTGAGAGAAATTTTCAG 7200

Qy 7201 CCGAGATAGATATCATCGGGTCCACAGACTGCTAGATGATGAGGGGTGTTTGTAT 7260
Db 7201 CCGAGATAGATATCATCGGGTCCACAGACTGCTAGATGATGAGGGGTGTTTGTAT 7260
Qy 7261 CCTAATGTTATCCCATGATGACGACAGACTTGTGTGCGAGTAAAGAGAGGTCAGGCTTC 7320
Db 7261 CCTAATGTTATCCCATGATGACGACAGACTTGTGTGCGAGTAAAGAGAGGTCAGGCTTC 7320
Qy 7321 AGATCAACAAAGACTGATTTCAATCTGATTTGAGGAGACCCCACTTTGATAGTGA 7380
Db 7321 AGATCAACAAAGACTGATTTCAATCTGATTTGAGGAGACCCCACTTTGATAGTGA 7380
Qy 7381 CTATATCTCTGAGCTCTGATCTCTCTTTTAAATGAGAGAGTAAATCCACATAG 7440
Db 7381 CTATATCTCTGAGCTCTGATCTCTCTTTTAAATGAGAGAGTAAATCCACATAG 7440
Qy 7441 CAGGCTGTGGGAGATTCAGAGATCAAAACAGCTGTGTATCAATCTGTGTTCTGTTC 7500
Db 7441 CAGGCTGTGGGAGATTCAGAGATCAAAACAGCTGTGTATCAATCTGTGTTCTGTTC 7500
Qy 7501 AGGCTATCAGACTGGGGTCTTGAAGTGAATTCACATCCAGCTTGGGTACAGAA 7560
Db 7501 AGGCTATCAGACTGGGGTCTTGAAGTGAATTCACATCCAGCTTGGGTACAGAA 7560
Qy 7561 CTGACACCAATCAACGAGCTGAGAGACTCTTGTCTTCAAGACAGACCTTGAAGCTTCA 7620
Db 7561 CTGACACCAATCAACGAGCTGAGAGACTCTTGTCTTCAAGACAGACCTTGAAGCTTCA 7620
Qy 7621 GGGCTGAGTGTGATCTGTTTCCCTTGTGTGCGCTGAACAGAAAGGGTGTGTCTGTGCTC 7680
Db 7621 GGGCTGAGTGTGATCTGTTTCCCTTGTGTGCGCTGAACAGAAAGGGTGTGTCTGTGCTC 7680
Qy 7681 CTGGGCTGCGCATGCGCAGAGAAAGCTGTCTCAATCTCAATCCCACTGTGTGCGGCTC 7740
Db 7681 CTGGGCTGCGCATGCGCAGAGAAAGCTGTCTCAATCTCAATCCCACTGTGTGCGGCTC 7740
Qy 7741 GACTTCTCTTCTTAAAGCGGCAATTAATGTTTGGCTTAAAGGCTCATCAATATCCG 7800
Db 7741 GACTTCTCTTCTTAAAGCGGCAATTAATGTTTGGCTTAAAGGCTCATCAATATCCG 7800
Qy 7801 CATCCCATCTCCAAATCTCAGCTGTGATGACCTTTTCCCTTAATGAGCTTAAGC 7860
Db 7801 CATCCCATCTCCAAATCTCAGCTGTGATGACCTTTTCCCTTAATGAGCTTAAGC 7860
Qy 7861 ATGCTGAGCGCATCAGACCGAGCGCTGTGCTCAATCTGTGAGCCATCTGTATACAC 7920
Db 7861 ATGCTGAGCGCATCAGACCGAGCGCTGTGCTCAATCTGTGAGCCATCTGTATACAC 7920
Qy 7921 TGCCGCGCGCCAGATACCTGTCAATGATGTGTCTGTGCTGTGCGGCTGTCTCCTG 7980
Db 7921 TGCCGCGCGCCAGATACCTGTGTCAATGATGTGTCTGTGCTGTGCGGCTGTCTCCTG 7980
Qy 7981 CTGCGAGATATCTGAGATGATGATTTCTGTGATCTTCTGTTTATGATGTGTCTGAT 8040
Db 7981 CTGCGAGATATCTGAGATGATGATTTCTGTGATCTTCTGTTTATGATGTGTCTGAT 8040
Qy 8041 TGGTGTAAAGTCAATTTCAATCAATGAGGCTGTGTTTATGATGTGTCTGCTC 8100
Db 8041 TGGTGTAAAGTCAATTTCAATCAATGAGGCTGTGTTTATGATGTGTCTGCTC 8100
Qy 8101 TGTGGTTCAGGCTGTGCTGTGTGATGAGATTTCTGTGTGATCCGGAAGATCCGCTG 8160
Db 8101 TGTGGTTCAGGCTGTGCTGTGTGATGAGATTTCTGTGTGATCCGGAAGATCCGCTG 8160
Qy 8161 ACCAGGCTGTACGAGCAATCTTCTCAACAGTGTGTCTTCTCTGTGTGCTGTGCTC 8220
Db 8161 ACCAGGCTGTACGAGCAATCTTCTCAACAGTGTGTCTTCTCTGTGTGCTGTGCTC 8220
Qy 8221 TTTGGCATTCAGTGGGCTGTGTTTCAAGATCCACTGTGATTTGAAGATCTTATTTGT 8280
Db 8221 TTTGGCATTCAGTGGGCTGTGTTTCAAGATCCACTGTGATTTGAAGATCTTATTTGT 8280

QY 8281 CATGTCATCTAGTATGTTTCCATTTTCCTGTCGGCTCTTAACAGAGTGCACACCCCATCATT 8340
 Db 8281 CATGTCATCTAGTATGTTTCCATTTTCCTGTCGGCTCTTAACAGAGTGCACACCCCATCATT 8340
 QY 8341 TACTTCTTGTGGGCTCCTTTAGGACGCGTCAAAATAGGACAACTGAAGCTGATTTC 8400
 Db 8341 TACTTCTTGTGGGCTCCTTTAGGACGCGTCAAAATAGGACAACTGAAGCTGATTTC 8400
 QY 8401 CAGAGGCTCTGACAGACACGCTGAGTGATGAGTGAAGGCTTCTCTACAGAA 8460
 Db 8401 CAGAGGCTCTGACAGACACGCTGAGTGATGAGTGAAGGCTTCTCTACAGAA 8460
 QY 8461 ACCCTGAGAGCTCTGAGAGAGATTTGAGAGAGAGAACTCTGCGCTGTACAGAC 8520
 Db 8461 ACCCTGAGAGCTCTGAGAGAGATTTGAGAGAGAGAACTCTGCGCTGTACAGAC 8520
 QY 8521 AGGACTTTGAGAGCAATGCTGCTGAGGTCCTGACACCTTGACAAATTATATGATTTTCTTAGCC 8580
 Db 8521 AGGACTTTGAGAGCAATGCTGCTGAGGTCCTGACACCTTGACAAATTATATGATTTTCTTAGCC 8580
 QY 8581 TTCTGCTCAGAAATGTCTCAGAGGTCCTGACAGGCTTCTCTACAGAA 8622
 Db 8581 TTCTGCTCAGAAATGTCTCAGAGGTCCTGACAGGCTTCTCTACAGAA 8622

RESULT 2

AB088140/c
 ID AB088140 standard; cDNA; 66494 BP.

XX AB088140;

AC 18-SEP-2002 (first entry)

DT 18-SEP-2002 (first entry)

XX Human osteoblast differentiation related cDNA SEQ ID NO 47.

DE Human osteoblast; stem cell differentiation; bone tissue deposition;

KW Human; osteoblast; stem cell differentiation; bone tissue deposition;

XX Human; osteoblast; stem cell differentiation; bone tissue deposition;

OS Homo sapiens.

XX WO200250301-A2.

PD 27-JUN-2002.

XX 18-DEC-2001; 2001WO-US048276.

PR 18-DEC-2000; 2000US-0255882P.

XX 24-APR-2001; 2001US-0285691P.

PA (GENE-) GENE LOGIC INC.

XX (PROC) PROCTER & GAMBLE CO.

PI Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;

XX Mertz L;

DR WPI; 2002-557663/59.

XX Use of genes and their expression profiles associated with osteoblast

PT differentiation for screening modulators bone formation, for diagnosing

PT or treating e.g. osteoporosis, or as markers for the differentiation

XX process.

XX Claim 1; SEQ ID NO 47; 78bp + Sequence Listing; English.

CC Involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
 CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
 CC or fibrous dysplasia. The present sequence is that of an osteoblast
 CC differentiation associated cDNA marker of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIDO
 CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 66494 BP; 17680 A; 13908 C; 14466 G; 20440 T; 0 U; 0 Other;

Query Match 16.2%; Score 1401; DB 6; Length 66494;

Best Local Similarity 77.7%; Pred. No. 0;

Matches 2034; Conservative 0; Mismatches 375; Indels 208; Gaps 20;

QY 15 GTCACTTATATCCAAACCGAAGAGATAT-----ACAAAGAAAGAACTATAGCC 68
 Db 64729 GTCACTTATATCCAAACCGAAGAGAGATATACAAAGAAAGAAAGAACTATAGCC 64670
 QY 69 AGTACCACTGATGATATATATCATGACAGAAATCCCAAGAAATATGCTAACCCATGCA 128
 Db 64669 AATATCCCTAATGAAATATGATGACAAATCTCTGACAAATATATGATGACAAATGCA 64610
 QY 129 ACGACATATCAAGAAATATATCCACATGTCAGTGGGTTTCATGACAGGGTG-CAAG 187
 Db 64609 ACGATATATCGAAAGATATATCCACATGTCAGTGGGTTTCATGACAGGGTGCA 64550
 QY 188 ATAGGTTATATCAATCAAGATGATATATATGATATATATATATATATATATATATAT 247
 Db 64549 ATAGGTTATATCAATCAAGATGATATATATGATATATATATATATATATATATAT 64490
 QY 248 AATACATGAT 307
 Db 64489 AATACATGAT 64430
 QY 308 TATGATTAATACCTTACAGAAATATGACATAGAAAGACATCTTATATATATATATAT 367
 Db 64429 TATGATTAATACCTTACAGAAATATGACATAGAAAGACATCTTATATATATATAT 64370
 QY 368 CATATATGACGAGACCAAGCAAGCAATATATATATATATATATATATATATATATAT 427
 Db 64369 CATATATGACGAGACCAAGCAAGCAATATATATATATATATATATATATATATAT 64311
 QY 428 CCTGAGAACTGGAACAGACAGAGATG-CTACTTACCACTTATATATATATATATAT 486
 Db 64310 TATAGAACTGGAACAGAGAAAGATGCGCACTTACCACTTATATATATATATATAT 64251
 QY 487 GGAAGTTTATGCGAGACATATGACAGAGAAAGAAATCAAGGCAACCAATATATATAT 546
 Db 64250 GGAAGTTTATGCGAGACATATGACAGAGAAAGAAATCAAGGCAACCAATATATAT 64191
 QY 547 AGAGAAAGTCAAGTGTCTGCTGATGATATATATATATATATATATATATATATAT 606
 Db 64190 AGAGAAAGTCAAGTGTCTGCTGATGATATATATATATATATATATATATATATAT 64131
 QY 607 AGACTCATCCAGAAAGCTCTAGAACTGATATATATATATATATATATATATATATAT 666
 Db 64130 AGACTCATCCAGAAAGCTCTAGATGATATATATATATATATATATATATATATAT 64071
 QY 667 ACTAAATGATCAAAATGATGACGTGATATATATATATATATATATATATATATAT 726
 Db 64070 AATCAATGATCAAAATGATGACGTGATATATATATATATATATATATATATATAT 64011
 QY 727 AATCAAGAACTCAAACTTATATATATATATATATATATATATATATATATATATAT 767
 Db 64010 AATCAAGAACTCAAACTTATATATATATATATATATATATATATATATATATATAT 63951
 QY 768 ACTTAAGAT 827
 Db 63950 TCTTGGAT 63891
 QY 828 AGCT---GACATCATGATGACACAAACAGTGAACACATCCCATGCTCATGATGG 883
 Db 63890 TGCTGAAGAAATATACAGGACACACAAATATATATATATATATATATATATATATAT 63831

QY 884 GTAGATCAATATTGTGAAAAATGACCAATTTGCCAAAAGCAATCTACAGTTCAATGCAA 943
 DB 63830 GTAGATCAATATTGTGAAAAATGACCAATTTGCCAAAAGCAATCTACAGTTCAATGCAA 63771
 QY 944 TTCCCAACCAAAATATCTCATCTCTTCTGACAGAACTGAAAGAAAAAACAATTTCTAAATTC 1003
 DB 63770 TGCCCAACCAAAATATCTCATCTCTTCTGACAGAACTGAAAGAAAAAACAATTTCTAAATTC 63711
 QY 1004 AT-----ATGAAACAAACCAACCAAAAAAATTTTCCGATGCCAAA 1050
 DB 63710 ATGAAATTTTAGAACCAATGCTTAAATCTGAAACCAAAAAAGAGCCCAATTTGCTAAA 63651
 QY 1051 GCAAGACTTAGCAAAAGAAACCAATCTGAGGCACTCATTTCCATCTTCAACCTATAC 1110
 DB 63650 GCAAG-CTAAGCAAAAGAAACCAATCTGAGG-ACACATTTGCTGACTTCAACCTATAC 63595
 QY 1111 TACAGGCTTAAATCAGCAAAACATCATGGCACTGACATTAATACTAGGCACTAGACCA 1170
 DB 63594 CA-TAGGCAATGCTATCATCAAAACATCATGGCTGATTAATAAC-AGGCACTAGACCA 63537
 QY 1171 TGGAAAAAGAGAGAAATCCAGAAATTAAGCAATTAATTAAGCACTGATTTTGAC 1230
 DB 63536 TGGAAAGAAATGAGAAACCAAAATTAAGCAAACTTAAGCACTGATTTTGAC 63477
 QY 1231 AAAGCAAAACAAAACATTAAGTGGGAAAGAAACATTTAGTTAAACAAATGCTGAGAT 1290
 DB 63476 AAAGCAAAACAAAACATTAAGTGGGAAAGAAACATTTAGTTAAACAAATGCTGAGAT 63417
 QY 1291 TATTGGCAACCAATGAGAAAGAAATGAACTGATCTCTGCTCTCACTTAATACAA 1350
 DB 63416 AATTGGCAACCAATGAGAAAGAAATTAACCTGAT-CTCATCTCTCACTGATTAATA 63358
 QY 1351 AATTGATACAGAGATGATCAAAAGACTTAAATCTGAGACTTAAACCACTAAATTTCTAGA 1410
 DB 63357 AATTGATACAGAGATGATCAAAAGACTTAAATCTGAGACTTAAACCACTAAATTTCTAGA 63298
 QY 1411 AGATTAACATGAGAAAAATGCTTCTAGACATTTCACTTGAAGCAAGACTTCATGGCAAGAA 1470
 DB 63297 AGATTAACATGAGAAAAATGCTTCTAGACATTTCTGAGCAAGAAAGTTCAAGCAAGAA 63238
 QY 1471 CCCAAAGTAAATGCAAAAAACAAAAATTAATGATAGACTTAATTAATTAATTAATTAATTA 1530
 DB 63237 CCCAAAGTAAATGCAAAAAACAAAAATTAATGATAGACTTAATTAATTAATTAATTAATTA 63178
 QY 1531 CTTTGGCAGCAAAACCAATCATTTAGCAGCAACCAACCAACCGAGTAGAGAA 1590
 DB 63177 C-TCTGCAAGCAAAACCAATCATTTAGCAGCAACCAACCAACCGAGTAGAGAA 63127
 QY 1591 AATCTTCAAAACCTAGCATCTGAATAAGCTTAATTCGGAATCCAGAGAACTCAA 1650
 DB 63126 -----TTTGGCAAACTTTCATCTGACAAAGCACTAATTCAGAACTTCAAAAGAAATCAA 63071
 QY 1651 ACAAAATGAGAAAGAAAGAAACCAATCCATGAAAGAGTGGGCTAAGGACATGAAATAG 1710
 DB 63070 ACAAAATGAGAAAGAAAGAAACCAATCCATGAAAGAGTGGGCTAAGGACATGAAATAG 63016
 QY 1711 ACAATTTCTAAAGAAATATATACAAATGAGCAACAAAC--AGAAAAAATCTTAAATC 1768
 DB 63015 ACAATTTCTAAAGAAATATATACAAATGAGCAACAAAC--AGAAAAAATCTTAAATC 62956
 QY 1769 ACTAATGATTAGGAAATGTAATCAACATGTAATGAGATACCACTTACTCTGCAAG 1828
 DB 62955 ACTAATGATTAGGAAATGTAATCAACATGTAATGAGATACCACTTACTCTGCAAG 62896
 QY 1829 AATGCTCATATTTTAAATTTAAATTAATGATGTTGGGTCTGTGTTGATTAAGG 1888
 DB 62895 AATGCTCATATTTTAAATTTAAATTAATGATGTTGGGTCTGTGTTGATTAAGG 62836
 QY 1889 AACACTTTTACAGCTGCTGGTGGGAAATGTAATCTTGGCAACCACTATGAGAAACAGTGTG 1948
 DB 62835 GACACTTTTACAGCTGCTGGTGGGAAATGTAATCTTGGCAACCACTATGAGAAACAGTGTG 62776

QY 1949 GAAATTTCTTAAGAACTAAAGTAAGTCAGACCTTTGATCCAGCAATCCCA----- 2000
 DB 62775 GAAATTTCTTAAGAACTAAAGTAAGTCAGACCTTTGATCCAGCAATCCCACTGCTGG 62716
 QY 2001 ----- 2000
 DB 62715 TATCTACCAAGCAAGAAAGATCATTTATACAAAAGATATCAATCAATGATGATTAAGAA 62656
 QY 2001 ----- 2000
 DB 62655 AATGATATTTTATATATATATAGTATACAGTATATACATATATATATATATATA 62596
 QY 2001 -----TTAATATGATTAATATATATATATATATATATATATATATATATATATATA 2057
 DB 62595 TA 62536
 QY 2058 AAGATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2117
 DB 62535 AAGATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 62476
 QY 2118 GGGGTAATCAGAAATGAAACCAACCAATCATATGTTCTCACTTCAAGTGGGCTTAA 2177
 DB 62475 GAGTAACTCAGAAATGAAACCAACCAATCATATGTTCTCACTTCAAGTGGGCTTAA 62416
 QY 2178 GCTGT---GAGGACAGCAAGGCAT-AGATGATTAATGAATCTGGGGACTTGAAGGGA 2233
 DB 62415 GCTGTGAGGAGGAGCGCAAGGCATTAAGCAATTAATGATGATGATGATGATGATGATGATGAT 62356
 QY 2234 AGATGAGAAAGAGGCGAGGAGTAAAGACTTACCAATGAGTACAGTGTCACTGCTCAG 2293
 DB 62355 AGGATGAGAGGAGTCAAGGAGTAAAGACTTACCAATGAGTACAGTGTCACTGCTCAG 62296
 QY 2294 GTGATGGTGCACCAAAATCTCAGAAATTAACCACTAAAGAACTTATCATGGAAGCAAC 2353
 DB 62295 CTGAGAGTGCACCAAAATCTCAGAAATTAACCACTTAAATGAACTTATCATGTAACCAAC 62236
 QY 2354 ACCACTGTTCCCAAAATCCCAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTTTC 2413
 DB 62235 ACCACTGTTCCCAAAATCCCAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTTTC 62177
 QY 2414 ACAGATTTAAAGAAAGTTCAGTGTGAGAGTTATTA 2450
 DB 62176 ATTGCATGAGGTAGGTTTATTTATCTCTGCTTACAA 62140
 RESULT 3
 ABA90193 0/c
 WP Sequence split into 5 fragments LOCUS ABA90193 Accession ABA90193
 WP Fragment Name Begin End
 WP ABA90193_0 1 11000
 WP ABA90193_1 100001 210000
 WP ABA90193_2 200001 310000
 WP ABA90193_3 300001 410000
 WP ABA90193_4 400001 465237
 ID ABA90193 standard; DNA; 465237 BP.
 XX
 AC ABA90193;
 XX
 DT 11-FEB-2002 (first entry)
 XX
 DB Human oestrogen receptor alpha gene.
 XX
 XX Human; oestrogen receptor alpha; ER; chromosome 6; Syme-2;
 KW synaptic nuclei expressed gene 2; haplotype; cytostatic; osteopathic;
 KW cardiant; vasotropic; gene therapy; vaccine; cancer; osteoporosis;
 KW cardiovascular disease; oestrogen receptor; ds.
 XX
 OS Homo sapiens.
 XX
 XX MO200162969-A2.
 XX
 XX 30-AUG-2001.
 XX

QY	1549	ATATCATTTAGCAGAGCAAAACAGACAAACCAACCGAGTGAGAGAAATCTTCAAAACTAAGC	1608
Db	5946	AACAGTCAGCAGATTAACACAGACAAACCAACGAATGGAGAAATCTTCAAACTTAATAC	5887
QY	1609	ATCTGACTAAGACCTAATATTCGGAAATCCAGAAAGAACTCAAAACAAATCAGCAAGAAAGAA	1668
Db	5886	ATCTGACAAAAGACTAATATCCAAAATCTCAATTAATCTCAAAACAAATCAGCA-----AT	5832
QY	1669	AGCAAAACAATCCCATGAAAGATGGGGCTTAAGACATGAATTGACAAATCTCAAAAAGAAAGAA	1728
Db	5831	TTTAAAAAATCCCTTCAGAAAAGTGGGCTTAAGACATGAATTGACAAATCTCAAAAAGAAAGAA	5772
QY	1729	TATACAAATGGCCCAACAA-----CAGAAAAAATGCTTAATCATCACTAATGATTAGGGAA	1784
Db	5771	TATACAAATGGCCCAAGAAACATTTCAAAAATAATGCTCAACATCACTAATGATTAGGGGCA	5712
QY	1785	ATGTAAATCAACACTGTATAGCGATCCAGCTTAATCTCTGCAAGAAATGGTCATTAATTTAA	1844
Db	5711	ATGTCAAAATCAAAACCGCAATGCCATATCCACTTACTCTGTAAAAACGGCCATATATCAAA	5652
QY	1845	AAATCTTAAAAATTAATAGATGTTGGTGGGTCTGTGTGATAAAGAAACCTTTACATCTGC	1904
Db	5651	AAATCTTAAAAACAGTAGATGTTGACGTGGAATGTGTATCAAGGGAACCTTCTCATTTGTC	5592
QY	1905	TGTGTGGGAATTAACCTTGCCCAACCACTATGAAAAACAGTGTGAAATTTCTTAAGGAA	1964
Db	5591	TGATGGGAATTTAACTATGTAACAAACCGCTATGAAAAAAGCGTGAAGATTCTTAAAGAA	5532
QY	1965	CTAAAAAGTAGATCCAGCACTTTGTATCAGCAATCCCAAT-----A	2003
Db	5531	CCAAAAGTAGAATCACTACCTTTGATCCAGCAATCCCACTAGGGTATCTAATCTCAGAGGAA	5472
QY	2004	AAATGTGTAATAATATAT-----	2021
Db	5471	AAATAGTCATTTATCAAAAAAAGATATCTTGCAACGCAATGTTTATATAGGGCAAAATTCACA	5412
QY	2022	-----	2021
Db	5411	ATTGCAAAATGCTGGAAACCAACCCAAATGCCCATCAATCATATGATGGAATAAAGTTCTG	5352
QY	2022	---TATTTATATACATGGAATACACCTCAGCCATAAAAAGATAAATGATGACATTTC	2078
Db	5351	TAGATATATATTTACTAGGGAATTACTACTCACACTATAAAAACAAGAAATTTAAACGGCATTC	5292
QY	2079	ACAGCAATCTGATGGAATTTGGAAGCCCTTAATTTCTAATGTGGGGTAACTCAGAAATGGAAT	2138
Db	5291	GCAATGACTGATGATGATTTGAAGCTATATTTCTAATGTAAGTAACTCAAGAAATGGAAT	5232
QY	2139	ACCAAAACATCATGTTCTCACTTAACAAGTGGGGCTTAAGCTGTGAGGACACGAAGCAT	2198
Db	5231	AACAAACATCATATGTTCTCACTGTTATATTTGGAGCTAAAGCTATGAGGATGCCAAAGCAT	5172
QY	2199	AGATGATATATATGAATCTTGGGGACTTGAAGGGGAAGGATGGAAGAGAGCGGAGGATTA	2258
Db	5171	AAAAATGATACAAAGAGATTTGGGGATTTGGAGGGGAAGGGTTGGAG-GGGGCAAGGAATTA	5113
QY	2259	AAGACTACACATGGGTTACAGTGTATCATCTGCTCAGGTATGAGGGTGCACAAATCTCGA	2318
Db	5112	AATATCTGAAATAGGGTGCAGTGTATCTGCTCAGGTATGAGGGTGCACAAATCTCGA	5053
QY	2319	AATTACCACTAAGAAGCTTATCAATGGAAGCAACCACTGTTCCCAAAATCCCAAT	2378
Db	5052	AATACCACTAAGAAGCTTATCTCATATTAACAAACACCACTGTATCCCAATTAACATATG	4993
QY	2379	GAAATTAATAATA 2390	
Db	4992	GAAAAAGAAAA 4981	

RESULT 4
AB087681_o/c
WP Sequence split into 5 fragments
LOCUS AB087681 Accession Ab087681

WP	Fragment Name	Begin	End
WP	ABQ87681_0	1	110000
WP	ABQ87681_1	100001	210000
WP	ABQ87681_2	200001	310000
WP	ABQ87681_3	300001	410000
WP	ABQ87681_4	400001	465237
ID	ABQ87681 standard; DNA; 465237 BP.		
XX	ABQ87681;		
AC			
DT	18-SEP-2002 (first entry)		
XX			
DE	Human oestrogen receptor alpha gene.		
XX			
KM	Human; oestrogen; receptor; oestrogen receptor alpha; cytostatic;		
KW	osteopathic; caduatic; cancer; osteoporosis; cardiovascular disorder		
XX	chromosome 6q25.1; gene; ds.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	variation	replace(18783,T)	
FT		/*tag= a	
FT		/strand name= "Single nucleotide polymorphism"	
FT		replace(18937,C)	
FT		/*tag= a	
FT	exon	/strand name= "Single nucleotide polymorphism"	
FT		18941. .19032	
FT		/*tag= b	
FT		/number= 1	
FT	intron	19033. .52817	
FT		/*tag= b	
FT		/number= 1	
FT		/cons_splice= (5'site:YES,3'site:NO)	
FT	variation	replace(19034,C)	
FT		/*tag= a	
FT		/strand name= "Single nucleotide polymorphism"	
FT	exon	52818. .52940	
FT		/*tag= b	
FT		/number= 2	
FT	variation	replace(52877,A)	
FT		/*tag= a	
FT		/strand name= "Single nucleotide polymorphism"	
FT	variation	replace(52901,A)	
FT		/*tag= a	
FT		/strand name= "Single nucleotide polymorphism"	
FT	intron	52941. .61149	
FT		/*tag= b	
FT		/number= 2	
FT		/cons_splice= (5'site:YES,3'site:NO)	
FT	exon	64150. .64280	
FT		/*tag= b	
FT		/number= 3	
FT	intron	64281. .166227	
FT		/*tag= b	
FT		/number= 3	
FT	variation	replace(64331,G)	
FT		/*tag= a	
FT		/strand name= "Single nucleotide polymorphism"	
FT	exon	166228. .166322	
FT		/*tag= b	
FT		/number= 4	
FT	intron	166323. .168001	
FT		/*tag= b	
FT		/number= 4	
FT	variation	replace(167950,G)	
FT		/*tag= a	
FT		/strand name= "Single nucleotide polymorphism"	
FT	variation	replace(167989,G)	
FT		/*tag= a	
FT		/strand name= "Single nucleotide polymorphism"	
FT	exon	168002. .168120	

FT		/+tag= b	
FT		/number= 5	
FT	variation	replace(168054,G)	
FT		/+tag= a	
FT		/standard name= "Single nucleotide polymorphism"	
FT	intron	168121..169542	
FT		/+tag= b	
FT		/number= 5	
FT	exon	/cons_splice= (5'site:YES,3'site:NO)	
FT		169543..169825	
FT		/+tag= b	
FT		/number= 6	
FT	variation	replace(169812,G)	
FT		/+tag= a	
FT		/standard name= "Single nucleotide polymorphism"	
FT	variation	replace(169823,G)	
FT		/+tag= a	
FT		/standard name= "Single nucleotide polymorphism"	
FT	intron	169826..169866	
FT		/+tag= b	
FT		/number= 6	
FT	exon	/cons_splice= (5'site:YES,3'site:NO)	
FT		169867..170678	
FT		/+tag= b	
FT		/number= 7	
FT	variation	replace(170035,A)	
FT		/+tag= a	
FT		/standard name= "Single nucleotide polymorphism"	
FT	variation	replace(170256,C)	
FT		/+tag= a	
FT		/standard name= "Single nucleotide polymorphism"	
FT	variation	replace(17038,G)	
FT		/+tag= a	
FT		/standard name= "Single nucleotide polymorphism"	
FT	variation	replace(170487,C)	
FT		/+tag= a	
FT		/standard name= "Single nucleotide polymorphism"	
FT	intron	170679..204911	
FT		/+tag= b	
FT		/number= 7	
FT	exon	204912..205102	
FT		/+tag= b	
FT		/number= 8	
FT	intron	205103..242969	
FT		/+tag= b	
FT		/number= 8	
FT	exon	242970..243086	
FT		/+tag= b	
FT		/number= 9	
FT	variation	replace(243055,T)	
FT		/+tag= a	
FT		/standard name= "Single nucleotide polymorphism"	
FT	intron	243087..306167	
FT		/+tag= b	
FT		/number= 9	
FT	variation	replace(243187,C)	
FT		/+tag= a	
FT		/standard name= "Single nucleotide polymorphism"	
FT	exon	306168..306503	
FT		/+tag= b	
FT		/number= 10	
FT	variation	replace(306292,A)	
FT		/+tag= a	
FT		/standard name= "Single nucleotide polymorphism"	
FT	variation	replace(306382,G)	
FT		/+tag= a	
FT		/standard name= "Single nucleotide polymorphism"	
FT	intron	306504..373639	
FT		/+tag= b	
FT		/number= 10	

FT	exon		3735640..373778
FT		/tag= b	
FT		/number= 11	
FT	intron	373779..422963	
FT		/tag= b	
FT		/number= 11	
FT	exon	422964..423097	
FT		/tag= b	
FT		/number= 12	
FT	variation	replace(423067,C)	
FT		/tag= a	
FT	intron	/standard_name= "Single nucleotide polymorphism"	
FT		423098..456533	
FT		/tag= b	
FT		/number= 12	
FT	variation	replace(423149,G)	
FT		/tag= a	
FT		/standard_name= "Single nucleotide polymorphism"	
FT	variation	replace(423163,G)	
FT		/tag= a	
FT		/standard_name= "Single nucleotide polymorphism"	
FT	variation	replace(423220,A)	
FT		/tag= a	
FT		/standard_name= "Single nucleotide polymorphism"	
FT	variation	replace(423232,G)	
FT		/tag= a	
FT		/standard_name= "Single nucleotide polymorphism"	
FT	variation	replace(423258,G)	
FT		/tag= a	
FT		/standard_name= "Single nucleotide polymorphism"	
FT	exon	456354..456537	
FT		/tag= b	
FT		/number= 13	
FT	intron	456538..460700	
FT		/tag= b	
FT		/number= 13	
FT	variation	replace(458706,C)	
FT		/tag= a	
FT		/standard_name= "Single nucleotide polymorphism"	
FT	variation	replace(459832,A)	
FT		/tag= a	
FT		/standard_name= "Single nucleotide polymorphism"	
FT	variation	replace(459913,G)	
FT		/tag= a	
FT		/standard_name= "Single nucleotide polymorphism"	
FT	variation	replace(460024,G)	
FT		/tag= a	
FT		/standard_name= "Single nucleotide polymorphism"	
FT	variation	replace(460056,T)	
FT		/tag= a	
FT		/standard_name= "Single nucleotide polymorphism"	
FT	variation	replace(460159,G)	
PT		/tag= a	
Query Match			
Best Local Similarity 77.8%; Pred. No. 1.6e-312;			
Matches 1970; Conservative 0; Mismatches 381; Indels 181; Gaps 17;			
QY	1	TGTATGAAGCCAAATCTCACTTTAATACCAAACAAGGATAATCAAAAAGAAAC	60
DB	7473	TCATAGAAGTACAGATTACTTAATATACCAAACAAGAAAGAACAT--AACCGAAAAAC	7417
QY	61	TATAGACCAGTACCACTGATGAATATACATGCAGAAATCCCAACAAAATATCTAGTTAAC	120
DB	7416	TACAGACCAATATCCTTGACGAACTAGATGCTTAAATCCTTAAACAAAATATCTAGCTTAC	7357
QY	121	CCAAATCCAACAGCATATCAAGAGATAATTCACCAATGTCAGTGGGTTTCATACCAAGG	180
DB	7356	TGAATCCAACAACATATCAAAAAGATTAATTCACCATGATCAAGTGGTTTCATACCAAGG	7297
QY	181	GTCGA-GGATAGGTTTAATATCACCAAGTCAATTAATGTGATACATCATATAACGAATT	239
DB	7296	ATGCAGGGATGGTTTAATATACACAGTCAATTAATGTGATACACCATATAACGAATT	7237

QY 240 AAAAAAATCATGATCATCTCATAGTGTGAAAAAGCAATTTGACAAAATCTAA 299
 DB 7236 AAAAAAATCATATATCATCTCAACAGATGAGAAAAAGCAATTTGACAAAATCTAG 7177
 QY 300 CATTTCTTATGATTTAAACCTTCAGCAAAAATCGACATAGAAAGACATACCTTAATGTA 359
 DB 7176 TATCCCTTATGATTTAAACCTTCAGCAAAAATCATCTTCAAGGACATCTCAATTA 7117
 QY 360 ATAAAGCCATATATGACGACCCACAGCAAACTTATATCTGATGAGGAAAAAGTTGAA 419
 DB 7116 ATAAAGCCATATATGACAAACCCACAGCAAACTATATGATGAGGAAAAAGTTGAA 7057
 QY 420 ACATTTCCCTGAGAACTGGAACCAAGAGAT- GCTACTTTCACCACTTCTATTCAC 478
 DB 7056 GGATTTCCCTGAGAACTGGAACCAAGAGATGCGCAATTCACCACTTCTTTCAC 6997
 QY 479 ATAGTATGGAAGTTTATGCGAGACATCAGACAGAGAAAGAAATCAAGGCAACCA 538
 DB 6996 ATAGTATGGAAGTTTCTAGTGAAGCAACAGAGAGAAAGAAATTAAGGCAATCCA 6937
 QY 539 ATCAATTAAGGAAGTCAAACTGTCTCTTCACTGATGATGATTTATCTAGAA 598
 DB 6936 ATTGACGAAGGAAGTCAAAATGTCTGCTGTTGCTGACAAATATGATTTTACTCAA 6877
 QY 599 AACCTTAAGACTATCCAGAAAGCTCTAGAACTGATATCAATTAATCAATTAAGTTCA 658
 DB 6876 AACCTTAAGACTCTCCAGAAAGCTCTTAAAGCTGATTAAGAAATTCAGAAAGTTTC- 6818
 QY 659 GGATACAAACTTAATGTGACAAATCAGTAGCACTGTATACCAAGAGTACCAAGCT 718
 DB 6817 GGATACAAAGTAAATGTGACAAATCAGTAGCTCTTCTATACCAAGAGTACCAAGCT 6758
 QY 719 GAGATCAAACTCAAGAACTCAACCTTTTACATAGCTGTAA-----AAATATC 769
 DB 6757 GAGATCAAACTCAAGAACTCAACCTTTTATGATAGCTGCAAAAGATGAATTAATATGC 6698
 QY 770 TTAAGATATTTCTTAACCAAGAGAGTG- AAGAACTCTACAGAGAAATCTCAAAACCA 828
 DB 6697 TTAAGATATATCTTAACCAAGAGAGTGAAAGAAATCTGACAGAGAAATCTCAAAACCT 6638
 QY 829 GCTGACATCATAGATGACAAACAGAGTGAACATCCCATGCTCATGAGGGGTGA 888
 DB 6637 GCTGAAAGAAATCATAGACGACAAATGAAACATCCATGCTCATGAGGGGTGA 6578
 QY 889 ATCAATATTTGAAATGACCAATATGCAAAAGCAATCTACAAATTCATGCAATCCC 948
 DB 6577 TTCAATATTTGAAATGAC-----TGCAAAAGCAATCTCAATTCATGCAATCCC 6523
 QY 949 ACCAAATATCATCATCTTTCTACAGAACTAGAAAAAAACATTTCTAAATTCATATG 1008
 DB 6522 ATTAATAT--ACCAACATTTCTACAGAACTTG- AAGACCAAAATTTCAAAATCTATAG 6467
 QY 1009 GAACAAACAAACCAAAAAAAGAAAAAAGCCGATAGCCAAAGCAAGCACTTAGCAAAAG 1068
 DB 6466 GAAC-----CCAAAAAAGCCCTGCTATAGCCAAAGCAAGCACTTAGCAAAAG 6421
 QY 1069 AACAAATCTGAGGACATCATTAACCATTTCTAACTATATCTCAAGGCTATATAC 1128
 DB 6420 AACAAATCTGAGGACATCATCTACAGATTTCAACATGATTTAAGGCCACGTAC 6361
 QY 1129 AAAACATCATGACCTGACATTAAGGCACTAGACCAATGAAAAAGAGAGAT 1188
 DB 6360 AAAACATCATGATCTGTATTAAGGCACTAGACCAATGAAACAGATGATGATAC 6301
 QY 1189 CCAAAATTAAGCCAAATATATATAGCACTGATTTTGAAGCAAGCAAAAGCACTA 1248
 DB 6300 CCAAAATTAAGCCAAATATATATAGCTGATGATTTTGAAGCAAGCAAAAGCACTA 6241
 QY 1249 AAGTGGGAAAAAGCATTTAGTTAACAAATGCTGAGATTAATTTGCAAGCCACATGT 1308
 DB 6240 AAGTGGGAAAAAGCACCTTTTGAACAAATGCTGAGATTAATTTGCTAGCCGAATGT 6181

QY 1309 GGAAGATGAATGAACTGGAATCCCTTGCTCTCACTTAATACAAAAATGATAGAGTGGAT 1368
 DB 6180 AGGAAGATGAATGGAAT- TCTCAATTTCTACCTTATCAAAAAATCACTTAGATGAT 6122
 QY 1369 CAAGACTTAATATGAGACCTTAAGCACTTAATTAATTTAGAGATTAATCAGAAAAAT 1428
 DB 6121 TAAGACTT-----TCGACCTGGAACCTTAATTAATTTAGAGATTAATCAGAAAAAT 6067
 QY 1429 GCTTCTGACATTCATAGGCAAGACTTATGAGCCAGAAACCAAAAGTAAATGCAAC 1488
 DB 6066 CTTTCTGACATTCAGGCTTAGGCAAGATTTCTATGACCAAGAACCAAAAGCAATTTCAAT 6007
 QY 1489 AAAAAAATTAATATGATGACCTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1548
 DB 6006 AAAAAAATTAATATGACCTGGAACCTTAATTAATTAATTAATTAATTAATTAATTAAT 5947
 QY 1549 AATCATTAGCAGACCAACAGACCAACCAAGAGTGAAGAAATCTTCAAACTAAGC 1608
 DB 5946 AATCATGAGAGATTAAGAGACCAACCAAGATGGAAGAAATCTTCAATCTATAC 5887
 QY 1609 ATCTGACTAGGACTAATATTCGGAATTCACAGAACTCAACAAATCAGCAAGAGAA 1668
 DB 5886 ATCTGACAAAGACTAATATTCGGAATTCACAAATTAATTAATTAATTAATTAATTAAT 5832
 QY 1669 AGCAAAATCCCATGAAAGAGTGGCTAAGAGATGAATGACAAATTTCTAAAAAGAA 1728
 DB 5831 TTTAAAAATCCCTTCAAGAAAGTGGCTAAGAGATGAATGACAAATTTCTAAAAAGAA 5772
 QY 1729 TATCAATGAGCAACAAA-----CAGGAAAAATGCTTAATCATCATATATAGGAA 1784
 DB 5771 TATCAATGAGCAACAAACCTTAACCAATTAATTAATTAATTAATTAATTAATTAAT 5712
 QY 1785 ATGTAATCAACACTGTAATGCAATACCACTTACCTCTGCAAGATGCTATATTTAA 1844
 DB 5711 ATGCAATCAACAAACGCAATCCATACCACTTACCTCTGTAAGAAAGGCAATATCAA 5652
 QY 1845 AATCTTAATTAATTAATGATGTTGGTGTCTGTGTGTATTAAGAAACATTTTACATG 1904
 DB 5651 AATCTTAATTAATTAATGATGTTGGTGTCTGTGTGTATTAAGAAACATTTTACATG 5592
 QY 1905 TGTGGAATGTAATCTTGCGCAACACTATGAAAAAGAGTGAATTTCTTAAGAA 1964
 DB 5591 TGTGGAATGTAATCTTGCGCAACACTATGAAAAAGAGTGAATTTCTTAAGAA 5532
 QY 1965 CTAAGATGATGCAATTTGATTCAGCAATCCAT-----A 2003
 DB 5531 CTAAGATGATGCAATTTGATTCAGCAATCCAT-----A 5472
 QY 2004 AATATGATTAATTAAT----- 2021
 DB 5471 AATATGATTAATTAAT----- 5412
 QY 2022 ----- 2021
 DB 5411 ATTGCAAAATCTGTGAAACCAACCAATGCCATCAATCAATGATGATTAAGATTTCTG 5352
 QY 2022 ----TATTAATTAATCAATGAAATCACTCAAGCCATTAATAAATTAATGATGATTC 2078
 DB 5351 TAGATATTAATTAATGAGATTAATCACTCAAGCCATTAATAAATTAATGATGATTC 5292
 QY 2079 ACAGCAATCTGATGAAATGAGACCTTATTTTAAGTGGGTGATCTCAGGAATGGA 2138
 DB 5291 GCAATGACCTGAGAGATGGAATGGAATTAATTTAAGTGAATGCAAGGAATGGA 5232
 QY 2139 ACCAAATCATATATGTTCTCACTTACAGAGTGGGGCTAAGCTGAGAGACAGAGGAT 2198
 DB 5231 AACAAATCATATATGTTCTCACTTATTAAGGCTAAGCTTATGAGATGCAAAAGAT 5172
 QY 2199 AGATGATTAATTAATGACTTGGGGAATTTGAGGGAAGATGGAAGAGAGGAGGAGATTA 2258
 DB 5171 AATATGATTAATTAATGAGATTTTGGGGAATTTGAGGGAAGGATTTGAG- GGGGAAAGATTA 5113
 QY 2259 AAGACTACAAATGAGGATGAGTGAACCTGCTCAGGTGATGAGGTGACCAAAATCTCAG 2318

```

Db      5112 AATACGCAATAGGTCAGCTGATCTGCTAGGTCATGATGGTCGACCAATCTCACA 5053
Qy      2319 AATTACCACTAAGAACTTATTCATGAGCAACCACTGTTCCCAAAATCCCAAT 2378
Db      5052 AATCACCACTTAAGAACTTATTCATGATTAACCAACCACTTACCCCAATTAACCTATG 4993
Qy      2379 GAATTAATAATA 2390
Db      4992 GAAAAAGAAAAA 4981

```

RESULT 5

ABX33717_0/c
 WP Sequence split into 5 fragments LOCUS ABX33717 Accession Abx33717

```

WP      Fragment Name      Begin      End
WP      ABX33717_0         1          110000
WP      ABX33717_1        100001      210000
WP      ABX33717_2        200001      310000
WP      ABX33717_3        300001      410000
WP      ABX33717_4        400001      465237
ID      ABX33717 standard; DNA; 465237 BP.

```

ABX33717;

26-FEB-2003 (first entry)

Gene encoding human oestrogen receptor alpha protein (ESR1).

Human; oestrogen receptor alpha; ESR1; cancer; osteoporosis;
 cardiovascular disorder; variant oestrogen receptor; ESR1 haplotype;
 ESR1 polymorphism detection; cytoskeletal; osteoparhbia; cardiac;
 chromosome 6q25.1; gene; single nucleotide polymorphism; SNP; ds.

Homo sapiens.

```

XX      Location/Qualifiers
XX      Key
XX      FT      variation
XX      replace(18783,T)
XX      /tag= ao
XX      /standard_name= "SNP"
XX      FT      variation
XX      replace(18937,C)
XX      /tag= ap
XX      /standard_name= "SNP"
XX      FT      exon
XX      18941..15032
XX      /tag= a
XX      /number= 1G
XX      FT      intron
XX      19033..52817
XX      /tag= b
XX      /number= 1G
XX      FT      variation
XX      replace(19034,C)
XX      /tag= aq
XX      /standard_name= "SNP"
XX      FT      exon
XX      52818..52940
XX      /tag= c
XX      /number= 1F
XX      FT      variation
XX      replace(52877,A)
XX      /tag= an
XX      /standard_name= "SNP"
XX      FT      variation
XX      replace(52901,A)
XX      /tag= am
XX      /standard_name= "SNP"
XX      FT      intron
XX      52941..64149
XX      /tag= d
XX      /number= 1F
XX      FT      exon
XX      64150..64280
XX      /tag= e
XX      /number= 1B
XX      FT      intron
XX      64281..166227
XX      /tag= f
XX      /number= 1B
XX      FT      variation
XX      replace(64331,G)
XX      /tag= al
XX      /standard_name= "SNP"

```

```

FT      exon
FT      166228..166322
FT      /tag= g
FT      /number= 1D
FT      intron
FT      166323..168001
FT      /tag= h
FT      /number= 1D
FT      variation
FT      replace(167950,G)
FT      /tag= ai
FT      /standard_name= "SNP"
FT      variation
FT      replace(167989,G)
FT      /tag= aj
FT      /standard_name= "SNP"
FT      exon
FT      168002..168120
FT      /tag= i
FT      /number= 1C
FT      variation
FT      replace(168054,G)
FT      /tag= ak
FT      /standard_name= "SNP"
FT      intron
FT      168121..169542
FT      /tag= j
FT      /number= 1C
FT      exon
FT      169543..169825
FT      /tag= k
FT      /number= 1B
FT      variation
FT      replace(169812,G)
FT      /tag= ag
FT      /standard_name= "SNP"
FT      variation
FT      replace(169823,G)
FT      /tag= ah
FT      /standard_name= "SNP"
FT      intron
FT      169826..169866
FT      /tag= l
FT      /number= 1B
FT      exon
FT      169867..170678
FT      /tag= m
FT      /number= 1A
FT      variation
FT      replace(170035,A)
FT      /tag= ab
FT      /standard_name= "SNP"
FT      variation
FT      replace(170068,T)
FT      /tag= ac
FT      /standard_name= "SNP"
FT      variation
FT      replace(170256,C)
FT      /tag= ad
FT      /standard_name= "SNP"
FT      variation
FT      replace(170368,G)
FT      /tag= ae
FT      /standard_name= "SNP"
FT      variation
FT      replace(170487,C)
FT      /tag= af
FT      /standard_name= "SNP"
FT      intron
FT      170679..204911
FT      /tag= n
FT      /number= 1A
FT      exon
FT      204912..205102
FT      /tag= o
FT      /number= 2
FT      intron
FT      205103..242969
FT      /tag= p
FT      /number= 2
FT      exon
FT      242970..243086
FT      /tag= q
FT      /number= 3
FT      variation
FT      replace(243055,T)
FT      /tag= ar
FT      /standard_name= "SNP"
FT      intron
FT      243087..306167
FT      /tag= r
FT      /number= 3
FT      variation
FT      replace(243187,C)
FT      /tag= ar
FT      /standard_name= "SNP"
FT      exon
FT      306168..306503

```

```

FT      /*tag= b
FT      /number= 4
FT      replace(306292,A)
FT      /*tag= at
FT      /standard_name= "SNP"
FT      replace(306382,G)
FT      /*tag= au
FT      /standard_name= "SNP"
FT      306504.373639
FT      /*tag= t
FT      /number= 4
FT      373640.373778
FT      /*tag= u
FT      /number= 5
FT      373779.422963
FT      /*tag= v
FT      /number= 5
FT      422964.423097
FT      /*tag= w
FT      /number= 6
FT      replace(423067,C)
FT      /*tag= av
FT      /standard_name= "SNP"
FT      replace(423163,G)
FT      /*tag= ax
FT      /standard_name= "SNP"
FT      replace(423220,A)
FT      /*tag= ay
FT      /standard_name= "SNP"
FT      replace(423232,G)
FT      /*tag= az
FT      /standard_name= "SNP"
FT      replace(423258,G)
FT      /*tag= ba
FT      /standard_name= "SNP"
FT      456354.456537
FT      /*tag= y
FT      /number= 7
FT      456538.460700
FT      /*tag= z
FT      /number= 7
FT      replace(459706,C)
FT      /*tag= db
FT      /standard_name= "SNP"
FT      replace(459832,A)
FT      /*tag= bc
FT      /standard_name= "SNP"
FT      replace(459913,G)
FT      /*tag= bd
FT      /standard_name= "SNP"
FT      replace(460024,G)
FT      /*tag= be
FT      /standard_name= "SNP"
FT      replace(460056,T)
FT      /*tag= bf
FT      /standard_name= "SNP"
FT      replace(460159,G)
FT      /*tag= bg
FT      /standard_name= "SNP"
FT      replace(460553,C)
FT      /*tag= bh
FT      /standard_name= "SNP"

Query Match      16.1%; Score 1390.4; DB 8; Length 110000;
Best Local Similarity 77.8%; Pred. No. 1.6e-312;
Matches 1970; Conservative 0; Mismatches 381; Indels 181; Gaps 17;

```

```

Qy      1 TGTATGAGCCGATGCTTAAATACCAAAACCGAAGAGATATACAAAAAGAAAC 60
Db      7473 TCTATGAGTCAAGCATTAATTAATACCAAAACCGAAGAGACT---AACGAAAAAC 7417
Qy      61 TATGAGCAGTACCACTGATGATATATACATGACAGAAATCCCAACAAATATAGCTAAC 120
Db      7416 TACAGACCAATATCCCTGACAGACATAGATCTAAATTCCTTAACAAATATAGCTAAC 7357
Qy      121 CCAATCCAGACATATCAGAGATATACCAACATTCGTAAGTGGTTTCATACACAGG 180
Db      7356 TGAATCCAAACAATATCAAAAAGATATTCACCATGATCAAGTGGTTTCATACACAGG 7297
Qy      181 GTGCA-GGATAGTTAATCATACAGATCAATTAATGATATCATCATTAACGAATT 239
Db      7296 ATGAGGATGTTAAATCATACAGATCAATTAATGATATCAACCATTAACGAATT 7237
Qy      240 AAAAAAATATCATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAATCTTA 299
Db      7236 AAAAAAATATCATATGATCATCTCAACAGATGCAAAAAAGCATTTGACAAATCTCA 7177
Qy      300 CATTCCTTATGATTAACCTTGACAAATCGACATAGAAAGACATATCATATGTA 359
Db      7176 TATCCCTTATGATTAACCTTGACAAATCGACATATCATATCAAGGACATATCTCAATTA 7117
Qy      360 ATAAAGCCATATATGACGACCCACAGCAACATTAATCTGATGGGAAAAAGTTGAA 419
Db      7116 ATAAAGCCATATATGACAAACCCACAGCAACATTAATCTGATGGGAAAAAGTTGAA 7057
Qy      420 ACATTTCCCTGAGAACTGGAACCAAGAT-GCTACTTTACCACTTCTATTTAAC 478
Db      7056 GGATTCCTCTGAGAACTGGAACCAAGATGGCCATTTCTCACTCTCTTTCAAC 6997
Qy      479 ATAGTAGGAAGTTTATGCGAGACATCAAGAGAAAGAAATCAAGGACCCAA 538
Db      6996 ATAGTAGGAAGTTTATGCGAGAAAGAAAGAAAGAAATCAAGGACCCAA 6937
Qy      539 ATCAATTAAGGAAGTCAAACTGCTGTTCACTGATGATATGATTTTATCTAGAA 598
Db      6936 ATGACGAAGGAAGTCAAACTGCTGTTCTGATGATGATTTGATTTCTCTCAA 6877
Qy      599 AACCTTAAGCTATCCAGAAAGCTCTAGAACGTATACATTAATTCAGTAAAGTTTCA 658
Db      6876 AACCTTAAGCTATCCAGAAAGCTCTAGAACGTATTAATTAATTCAGTAAAGTTTCA 6818
Qy      659 GGATACAACTAAATGTACAAATCAATGATGACCTGTATACCAACAGTACCAAGCT 718
Db      6817 GGATACAACTAAATGTATGACAAATCAATGATGACCTGTATACCAACAGTACCAAGCT 6758
Qy      719 GAGATCAATCAAGAACTCAACACTTTTACATAGCTGTAA-----AAATAC 769
Db      6757 GAGATCAATCAAGAACTCAACCTCTTTATGATGATGCTGCAAAAGATGAATTAATAATGC 6698
Qy      770 TTAAGATATCTTAACCAAGAGAGTG-AAAGACTCTACAGAGAAACTCAAAACCA 828
Db      6697 TTAAGATATCTTAACCAAGAGAGTAAAGAAATGTGACAGAGAAACTCAAAACCACT 6638
Qy      829 GCTGACATCATAGATGACAAACAGATGGAACACATCCCATCTCATGATGGGTAGA 888
Db      6637 GCTGAAAGAAATCATAGACACACAAATGGAACACATCCCATCTCATGATGGGTAGA 6578
Qy      889 ATCAATATGTAAGTACCATATTTGCAAAAGCAATCTACAGTTCAATGCAATTTCC 948
Db      6577 TTCAATATTTGAAATGAC-----TGCAAAACCAATCTCAATTCATCAATGCC 6523
Qy      949 ACCAAATATCATATCTTCTGACAACTTAAGAAAAACATTTCTTAATTCATATG 1008
Db      6522 ATTAATAT---ACCAATTTCTTACAGAACTTG-AGACCAATTTCTTAATTCATATG 6467
Qy      1009 GAACAAACCAAAAAAAGAAACCCGATAGCCAAAGCAAGCATTTAGCAAAAG 1068
Db      6466 GAAC-----CCAAAGAAAGCTGATAGCAAGCAAGCATTTAGCAAAAG 6421
Qy      1069 AACAAATGAGAGCATCATTAACCATCTTCAAACTATATCAAGGCTATATATACC 1128

```

```

Db      6420 AACAAATCTGGAGGCACTACCAAGATTCAAACTGATATTAAGGCCACAGCTACC 6361
Qy      1129 AAAACATCATGGCACTGACATTAATACTAGGCACTAGACCAATGAAAAAGAGAAAT 1188
Db      6360 AAAACAGCATGCTAGTATTAATAATAGGCACTAGACCAATGAAACAGAAATGATTAAC 6301
Qy      1189 CCAAGAAATTAAGCCAAATTAATTAAGCCAACTGATTTTGAACAACCAAAACAAATTAAT 1248
Db      6300 CCAGAAATTAAGCCAAATTAATTAAGCCAACTGATTTTGAACAACCAAAACAAATTAAT 6241
Qy      1249 AAGTGGGAAAAAGCACTTCTAGTAAACAATGGTGTGCTAGATTTATTTGGCAAGCCATCT 1308
Db      6240 AAGTGGGAAAAAGCACTTCTAGTAAACAATGGTGTGCTAGATTTATTTGGCAAGCCATCT 6181
Qy      1309 GGAAGAAATGAACCTGGATCCCTGTCTCTCACTTAATTAACAAAATTTGATCAAGATGAT 1368
Db      6180 AGGAGAAATGAACCTGGAT-TCTCATTTCTCACTTAATTAACAAAATTTGATCAAGATGAT 6122
Qy      1369 CAAGAATTAATCTGAGACCTTAACCAATTAATTTCTGAAGATTAATCAAGAAAT 1428
Db      6121 TAAAGACTT-----TCGACCTGAAACTCTAAAAATTTGAGAGATTAACCTGAAAAAC 6067
Qy      1429 GCTTCTAGACATTCCTTAGGCAAGAACTTCAATGGCCAAAGCCAAAGTAAATGCAAC 1488
Db      6066 CTTTCTAGACATTTGGCTTAGGCAAGAACTTCAATGGCCAAAGCCAAAGTAAATGCAAC 6007
Qy      1489 AAAAAAATAATTAATGATAGGACTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1548
Db      6006 AAAAAAATAATTAATGATAGGACTTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1547
Qy      1549 AATCATTAAGCAAGCAACGACCAACCCAGTGTAGAGAAATTTTCACAAACTTAAC 1608
Db      5946 AACACTGACGAAAGTAAACAGAACCCAGATGAGAGAAATTTTCACAAACTTAAC 1587
Qy      1609 ATCTGACTAAGCACTTAATTCGGAATTCACAGAAGCACTCAAACTCAAGAAAGAA 1668
Db      5886 ATCTGACAAAGCACTTAATTCGGAATTCACAGAAGCACTCAAACTCAAGAAAGAA 1667
Qy      1669 AGCAAAACAATCCATGAAAGAGTGGGCTAAGACATGAATTAACAATTTCAAAAGAA 1728
Db      5831 TTTAAAAATCCCTTCAAGAAAGTGGGCTAAGACATGAATTAACAATTTCAAAAGAA 1727
Qy      1729 TATTAACAATGGCCAAACAA-----CAGAAAAAATGCTTAATCACTTAATGATTAAGGA 1784
Db      5771 TATTAACAATGGCCAAACAA-----CAGAAAAAATGCTTAATCACTTAATGATTAAGGA 1783
Qy      1785 ATGTAATCAACACTGTAATGCGATACCACTTACTCTGCAAGAAATGTCATTAATTTAA 1844
Db      5711 ATGCAATTAACCAAGCAATGCTACCACTTACTCTGTAAGAAAGCCCAATTAATCAAA 1843
Qy      1845 AAATCTAAAAATTAATGATGTTGGTGGTCTGTGTGATTAAGAAACACTTTTACACTGC 1904
Db      5651 AAATCTAAAAATTAATGATGTTGGTGGTCTGTGTGATTAAGAAACACTTTTACACTGC 1903
Qy      1905 TGGTGGGAAATGTAATTTGCGCAACCACTTAATGAAAAAGATGGAATTTCTTAAGAA 1964
Db      5591 TGAATGGAATGTAATTTGCGCAACCACTTAATGAAAAAGATGGAATTTCTTAAGAA 1963
Qy      1965 CTAAAGTATGATGCACTTTGATTCAGCAATCCCAAT-----A 2003
Db      5531 CCAAAAGTATGATGCACTTTGATTCAGCAATCCCAAT-----A 2002
Qy      2004 AATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2021
Db      5471 AATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2020
Qy      2022 ----- 2021
Db      5411 ATTGCAAAATCTGGAACCAACCAATGCCATCAATGAATGATTAAGATTTCTG 5352
Qy      2022 ---TATTTATATACATGGAATCAACTGACATTAATAAGAAATTAATGATGACATTC 2078

```

```

Db      5351 TAGTATATATATAGGAGTACTACTGACTATATAAAGAAAGAAATTAACGGCAATTC 5292
Qy      2079 ACAGCAATTAATGATGAAATTTGAGACCCCTTATTTCTAAGTGGGTAATCTCAGAAATGAAA 2138
Db      5291 GCAATGACCTGATGATGATTTGAGACCTTATTTCTAAGTGGGTAATCTCAGAAATGAAA 5232
Qy      2139 ACCAATCATATATGTTCTCACTTCAAGTGGGCTAAGCTGTGAGACAGAGCAT 2198
Db      5231 AACAAATCATATATGTTCTCACTTCAAGTGGGCTAAGCTGTGAGACAGAGCAT 5172
Qy      2199 AGAATGATTAATGAATCTGTGGGACTTGAAGGGAAGATGGAAGAGGCGAGGATTA 2258
Db      5171 AAATGATTAATGAAGGATTTGGGAAATTTGAGAGGAAAGGTTGAG-GGGGCAAGGAATTA 5113
Qy      2259 AAGCTACCAATGGGTAATGATGACCTGCTCAGGTGATGGGTGACCAAAATCTCA 2318
Db      5112 AATACGCAAAATGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 5053
Qy      2319 AATTACCAATTAAGAACTTATCCATGGAAGCAACCACTGTTCCCAAAATCCCAAT 2378
Db      5052 AATACCACTTAAGAACTTATCTCATTAATTAACCAACCACTGTTACCCCAATCACTATG 4993
Qy      2379 GAAATTAATA 2390
Db      4992 GAAAAAGAAAA 4981

RESULT 6
ABX08336_13
Continuation (14 of 17) of ABX08336 from base 1300001 (Human phosphodiesterase 4D (PDE4D
WP Sequence split into 17 fragments LOCUS ABX08336 Accession Abx08336
WP Fragment Name Begin End
WP ABX08336_00 1 110000
WP ABX08336_01 100001 210000
WP ABX08336_02 200001 310000
WP ABX08336_03 300001 410000
WP ABX08336_04 400001 510000
WP ABX08336_05 500001 610000
WP ABX08336_06 600001 710000
WP ABX08336_07 700001 810000
WP ABX08336_08 800001 910000
WP ABX08336_09 900001 1010000
WP ABX08336_10 1000001 1110000
WP ABX08336_11 1100001 1210000
WP ABX08336_12 1200001 1310000
WP ABX08336_13 1300001 1410000
WP ABX08336_14 1400001 1510000
WP ABX08336_15 1500001 1610000
WP ABX08336_16 1600001 1691080

Query Match 16.0%; Score 1375.6; DB 6; Length 110000;
Best Local Similarity 77.1%; Pred. No. 4.7e-309; Indels 182; Gaps 19;
Matches 1993; Conservative 0; Mismatches 409;

1 TGTATGAGCAATGTCACCTTAATACCAAAACGAGAAAGATTA--CAAAAAAGAA 58
25493 TCTATGAGCCAGACATCAACCTTAATGTAACCAAAACGAGAAAGATTAACCAAAAAAGAA 25552
59 ACTATGAGCCAGTACCACTGATGAATATATCATGCAAGAAATCCCAAAATACTAGCTTA 118
25553 ACTATGAGCCAGTACCACTGATGAATATATCATGCAAGAAATCCCAAAATACTAGCTTA 25612
119 ACCCAATCCCAAGCACTATCAAGAAATATATCATGCAAGAAATCCCAAAATACTAGCTTA 178
25613 ACTGAAATCCCAAGCACTATCAAGAAATATATCATGCAAGAAATCCCAAAATACTAGCTTA 25672
179 GGGTGCAG-GATAGTTAATCATCAAGTCAATTAATGATGATCAATTAACAGAA 237
25673 GGGTGCAGAGATGTTTAAATGTAACAGTCAATTAATGATGATCAATTAACAGAA 25732
238 TTAATAAACAATAATCAATGATCATCTCAATGATGCTGAAGAAAGATTTGAACAAATCT 297
25733 TTAATAAACAATAATCAATGATCATCTCAATGATGCTGAAGAAAGATTTGAACAAATCT 25792

```

QY 298 AACATTCTTTAGATTTAAAACTTCAGCAAAATCGACATGAGAGACATACCTTAATG 357
 DB 25793 AGCATCCCTTTATGATTTAAAGCTCTCAGCAAAATCAGATACAGGGGACATACATTAATG 25852
 QY 358 TAAATTAAGCATATATGACGGACCCACAGCAAACTTATATCTGAATGGGAAAAAGTTGA 417
 DB 25853 TAAATTAAGCATATATGACCAACCCACAGCAAACTTATATCTGAATGGGAAAAAGTTGA 25912
 QY 418 AAACATTGCTCCCTGAGAACTGGAACAAGACAGATG-CTACTTTACCACTTCTATTTCA 476
 DB 25913 AAGAAATTCCTCTGAGAACTGGAACAGACATATGATGCCACTCTCACCACTTCTTTCA 25972
 QY 477 ACATAGTAGTGAAGTTTTCAGAGCAATCAGACAGAGAAAGAAATTCAGAGGCAACC 536
 DB 25973 ACATAGTAGTGAAGTCTTACAGAGCAATCAGACAGAGAGGAGAAATTAAGGCAATCC 26032
 QY 537 AAATCAATTAAGAGAGAGTCAAACTGTCCCTGTTCACTGATGATATGATTTATACCTAG 596
 DB 26033 AAATCGGTAAAGAGAGTCAAACTGTCTCTGTTGCTGATGATATGATTTATACCTAG 26092
 QY 597 AAAACCTTAAGAGTCTATCCAGAAAGCTCCTAGAACTGATACATTAATTCAGTAAAGTTT 656
 DB 26093 AAAACTTAAGAACTCTCCAGCAAGCTCTAGAACTGATTAATGAAATTCAGAAAGTTT 26152
 QY 657 CAGGATACAACTAATATGTACACAAATCAGTACGCTATACCAAGTACGACCAAG 716
 DB 26153 CTGATATCAAGATTAATGTACAAATCAGTACGCTTCTATATACCAAGTACCAAG 26212
 QY 717 CTGATATCAAGATTAAGAACTCAAACTTTTACATAGCTGTAA-----AAA 765
 DB 26213 GGGGAATCAAAATCAAGAACTCAAACTTTTACATAGCTGTAA-----TAAATTA 26272
 QY 766 ATACTTAAGATATTTCTTACCAAGAGTGAAGAGCTCTACAGAAATCTACAAATC 825
 DB 26273 ATACTTAAGATATATCTTACAAAGAGTGAAGAGCTCTTACAGAAATCTACAAATC 26332
 QY 826 ACAGCTGA-----CATCATAGTACACAAACAGTGAACACATCCCATGCTCAGTAT 881
 DB 26333 ACTCTGAAGAGATTCATAGACAAATACAAATAGGACACATCCGATATCATGAT 26392
 QY 882 GGGTGAATCAATATTTGGAAGAAATGACATATTTGCCAAAGCAATCTTCAAGTTCAATGC 941
 DB 26393 GGGTGAATCAATATTTGGAAGAAATGACATATCTGCCAAAGCAATCTTCAAGTTCAATGC 26452
 QY 942 AATCCCAACCAAAATATCATCATCTTCTTACAGAACTTGAAGAAATCAATTTCAAAAT 1001
 DB 26453 AATCCCAACCAAAATATCATCATCTTCTTACAGAAATTTAG-AAAAACCAATTTCTAAATC 26511
 QY 1002 TCATATGAGAACAAACCAAAAAAAAAAAAAAAAAAAAAAATCCGCAATGCAAGCAAGCTTAG 1061
 DB 26512 TCATATGAGAAC-----CAAAAAAAAAAGCCGCAATGCAAGCAAGCAAGCTTAG 26557
 QY 1062 CAAAAAGAACAAATCTGAGAGCATCACTTACCATCTTCAAGATCTTCAAGGCTAT 1121
 DB 26558 CAAAAAGAACAAATCTGAGAGCATCACTTACCATCTTCAAGATCTTCAAGGCTAT 26617
 QY 1122 AATCAACCAAAACATCTGAGCACTGACATTAATAGGACATAGACCAATGGAAGAAAG 1181
 DB 26618 AGTACCAAAATGAGCATGCTGATTAATTAATGACATATGACCAATGGAAGCAAGAT 26677
 QY 1182 AGAGATTCAGAAATTAAGCCAAATATTAATGACCAATGATTTTGAAGCAAGCAAA 1241
 DB 26678 AGAGATTCAGAAATTAAGCCAAATATTAATGACCAATGATTTTGAAGCAAGCAAA 26737
 QY 1242 AAAATTAAGTGGGAGAAA-ACATTCATGTTAACTAATGCTGAGATTAATTTGGCAG 1300
 DB 26738 AAAATTAAGTGGGAGAAAAGATTAACCTTTCAACAAATGCTGAGATTAATTTGGCAG 26797
 QY 1301 CCACATGTGAGAGATGAATGAACTGATCCCTTGTCTCTCACTTAATCAAAATTTGATACA 1360
 DB 26798 CCACATGTGAGAGATGAATGAACTGATCC--TATCTCTCACCTGATTAACAAATTCACCTCA 26855

QY 1361 AGATGATCAAAAGCTTAAATCTGAGACCTTAAACCATTAATTTTGAAGATTAATCTC 1420
 DB 26856 AGATGATTAAGGGCTTAAACCTTAAAGCTGAAATCTAT-GAAATTTTGAAGATTAATCTT 26914
 QY 1421 AGAAAAATGCTTCAAGCACTTCACTTAGGCAAGAACTTCAAGGCAAGAACCCAAAGTA 1480
 DB 26915 GAAAAATCCCTTCAAGCACTTGGCTTGAAGCAAGATTTCAATGACCAAGAACCCAAAGTA 26974
 QY 1481 AATGCAACAAAAACAAAAATTAATAGATAGGACTTAAATTAAC-TAAAAAGCTTTTGGCC 1539
 DB 26975 AATGCAATTAACAAAGATTAATAGCTGGGACCTCATTAATCTTACGACTTTTGGAG 27034
 QY 1540 AGCAAAAACATCAATGACAGCAAAACAGCAAAACCCACGAGTGAAG-AAAATCTTCA 1598
 DB 27035 GGCAAAAGGAAACAGTCAAGAGTGAACAGCAAAACCCACAGAGTGGGAAAAATCTTCA 27094
 QY 1599 CAATCTAAGCATCTGACCTAAGGACTTAATATCCGGAATTCACAAAGAACTCAACAAATCA 1658
 DB 27095 CAATCTAAGCATCTGACCAAGG-GTAGTATCAAGATTCACAGAGACCCCAACAAATCA 27153
 QY 1659 GCAAGAGAAAGCAAAACATCCATGAAGAGTGGGCTAAGGACATGATAGCAATTTCT 1718
 DB 27154 GTAAGAAAAACAAACAAATCCATCAAAAAGTAAAGGCAATGATGAGCAATTTCA 27213
 QY 1719 CAAAAAGATATCAAAATGGCCAAACAAAC--AGAAAAATGCTTAACTATCATATGA 1776
 DB 27214 CAAAAAGATATCAAAATGGCCAAACAAACATATGAAAAATGCTCAACATCATATGA 27273
 QY 1777 TTAGGAAATGTAAATCAAC-CTGTAATCCGATACACCTTACTCCTGCAAGATGT 1834
 DB 27274 TCAAGAAATGCAATCAAAACAAACAAATGTATACCAAGCTTCTTGAAGATGGC 27333
 QY 1835 CATATTTTAAAAATCT--AAAAATATAGATGTTGGTGGCTGTGATTAAGAAACA 1892
 DB 27334 CATATTTTAAAAATTTTAAAAACAGTATGTTGGATGGAAGGGGATGACAGAAACA 27393
 QY 1893 CTTTACATGCTGCTGGGAAATGTAACTTGGCAACCATATGGAAGAAACAGTGTGAAA 1952
 DB 27394 CTTTACATGCTGCTGGGAAATGTAACTTGGCAACCATATGGAAGAAACAGTGTGAAA 27453
 QY 1953 TTTCTTAAGGAACCTAAGATGATGACCAATTTGATCAGCAATCCAT----- 2001
 DB 27454 TTCTTAAAGAACTTAAAGTGAACCTACCTTATCAGAGTCCACTATAGATATC 27513
 QY 2002 ----- 2001
 DB 27514 TACCAAGAGAAAGAAAGTCAATTTTGAAGAAACACTTGTACAGATTTATAGCA 27573
 QY 2002 ----- 2001
 DB 27574 GCACAAATTCAAATTTGCAAACTGTGAACCTAACCAATGTCCATCATATTAAGTG 27633
 QY 2002 -TAAATATGATTAATATATATTTATTTATTCATGAAATCACTCAGCATTAAGAAAG 2060
 DB 27634 ATTAAGAAACCTGTGTGATCAGATATATATCAATGAATATCTACACAGCTATGAAGG 27693
 QY 2061 AATAAAAATGATGACATTCACAGCAATCTAGATGGAATTTGAGACCTTATTTAAGTGG 2120
 DB 27694 AATGAATTAACAGATTTTGAAGTGAAGCTTGAAGATTTAATTTCTAAGTAA 27753
 QY 2121 GTAATCTGAGATGGAAGAAACCAACATCATATGTTCTCATCTTACAGATGGGGCTAAGCT 2180
 DB 27754 GTAATCTGAGATGGAAGAAACCAACATCATATGTTCTCATCTGATATGAGATCTAAGCT 27813
 QY 2181 GTGAGGACAGGAAGCAAT-AGATGATATATGAACTCTGGGGGCTTGAAGGGGAAGATG 2239
 DB 27814 ATGAGGACACAAAGATATATGAATGATATCAATGACCTTTGGGAGCTTGGGGGAAGATG 27873
 QY 2240 GAAAGAGGCGAGGATTAAGATCTACCAATGAGTACAGTGTACATGCTCAGGTGATG 2299
 DB 27874 GAGAGGGGTGAGGATTAAGATTTACAAATATGATGAGTATATCTGCTTGGGTATG 27933
 QY 2300 GGTGACCAAAATCTCAGAAATTAACACTAAGAACTTATTCATGAGAAACA--ACACA 2357

QY 1599 CAACTAAGCATCTGACCTAAGCATTAATTCGGGAATCCACAGGAATCTCAACAAATCA 1658
 Db 27154 CAATCTAATACCTCGACAAAGG-GTAGATATCCAGATCTACAAAGAACCCCAACAAATCA 27212
 QY 1659 GCAAGAGAAAGCAAAACAATCCCATGAAGAGTGGGCTAAGGACATGAATGACAATCT 1718
 Db 27213 GTAAAGAAAAAACAACATCCCATCAAAAGTAGGCTTAAGGGCATGAGTGGCAATTC 27272
 QY 1719 CAAAGAGATATACAAATGCGCAACAAAC--AGAAAAATGCTTAAATCATCATTAATGA 1776
 Db 27273 CAAAGAGATATACAAATGCGCAACAAACATATGAAAAAATGCTCAATCATCATTAATGA 27332
 QY 1777 TTAGGAAATGTAATCAACA--CTGTATGCGATACCACTTACTCTCGAAGATGCT 1834
 Db 27333 TCAAGGAATCAATCAAAACACAAATGTAATACCAACCGTACTCTCGAAGATGCT 27392
 QY 1835 CATATTTTAAATCT--AAAAATATAGATGTTGGTGGTCTGTGTGATTAAGGAACA 1892
 Db 27393 CATATTAATAAATTTTAAAAAACAATGATGTTGGCATGAAAGGGGTGATCAGGAACA 27452
 QY 1893 CTTTACACTGCTGGTGGAAATGTAACCTTCGCAACCACTATGGAACAGTGTGAAA 1952
 Db 27453 CTTCTACGCTGCTGGTGGAAATGCAACTAGTACAGCAATTAATGGAACAGTGTGGA 27512
 QY 1953 TTTCTTAAAGACTTAAAGTATGATGACCACTTGTATCAGCAATCCAT----- 2001
 Db 27513 TTCTTAAAGACTTAAAGTATGATGACCACTTGTATCAGCAATCCATCTAGATATC 27572
 QY 2002 ----- 2001
 Db 27573 TACCAAGAGAAAGAACTATTTTGAAGAAACACTTGTACAGTATGTTATAGCA 27632
 QY 2002 ----- 2001
 Db 27633 GCACAAATTCACAAATGCAAACTGTGAACTTAACCAATGTCATCATTAATGTCG 27692
 QY 2002 -TAAATATGTAATATATATATTTATATATTCATGAAATACAACTCAGCCATAAAAAG 2060
 Db 27693 ATTAAGAACTGTGTGATACATATATATATCAATGAATATCTACACAGCTATGAAAAAG 27752
 QY 2061 AATAAATGATGATTCACAGCAATCTAGATGGAATGGAAGCCCTTAATCTAAGTGGG 2120
 Db 27753 AATGAATTAACAGCAATTCAGTGCCTGATGAGATGGAAGACTATTAATCTAAGTGA 27812
 QY 2121 GTAACCTAGGAATGAAAAACCAACATCATATGTTCTCACTTAAGAGTGGGGCTAAGCT 2180
 Db 27813 GTAACCTAGGAATGAAAAACCAACATCATATGTTCTCACTGATATGAGGATCTAAGCT 27872
 QY 2181 GTGAGACACGAGGACAT-AGATGATATTAATGAATCTTGGGGACTTGAAGGGAAGATG 2239
 Db 27873 ATGAGGACACAAATATTAAGATATGAATGAATGAATGAACTTGGGGGAAGAGTGG 27932
 QY 2240 GAAGAGGCGAGGATTAAGAACTAACAATGCGTACAGTGTACATGCTCAAGTGTATG 2299
 Db 27933 GAGAGGGGGTGAAGGATTAAGAACTAACAATATGCTGAGTGTATGCTGTGGGTATG 27992
 QY 2300 GGTCACCAAAATCTCAGAAATTAACCACTAAAGAACTTATCCATGGAAGCA--ACACCA 2357
 Db 27993 GGTCACCAAAATCTCAGAAATTAACCACTAAGAACTTATTAATTAATTAATTAACCA 28052
 QY 2358 CCGTGTCCCAAAATCCCAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2417
 Db 28053 CCGTGTCCCAAAATCCCAATGAATTAATTAATTAATTAATTAATTAATTAATTAATTA 28112
 QY 2418 AATT 2421
 Db 28113 GTTT 28116

RESULT 8
 ADN97989 13
 Continuation (14 of 17) of ADN97989 from base 1300001 (Human phosphodiesterase 4D genome)

WP Sequence split into 17 fragments LOCUS ADN97989 Accession Adn97989
 WP Fragment Name Begin End
 WP ADN97989_00 1 110000
 WP ADN97989_01 100001 210000
 WP ADN97989_02 200001 310000
 WP ADN97989_03 300001 410000
 WP ADN97989_04 400001 510000
 WP ADN97989_05 500001 610000
 WP ADN97989_06 600001 710000
 WP ADN97989_07 700001 810000
 WP ADN97989_08 800001 910000
 WP ADN97989_09 900001 1010000
 WP ADN97989_10 1000001 1110000
 WP ADN97989_11 1100001 1210000
 WP ADN97989_12 1200001 1310000
 WP ADN97989_13 1300001 1410000
 WP ADN97989_14 1400001 1510000
 WP ADN97989_15 1500001 1610000
 WP ADN97989_16 1600001 1691138

Query Match 16.0%; Score 1375.6; DB 12; Length 110000;
 Best Local Similarity 77.1%; Pred. No. 4.7e-309;
 Matches 1993; Conservative 0; Mismatches 409; Indels 182; Gaps 19;

QY 1 TGTATGAAGCCATGATCTTAATTAATCAAAACAGAGAAAGATATA--CAAAAAAGAAA 58
 Db 25551 TCTATGAAGCCAGATCATCCCTAGTACCAAAACAGAGAAAGATATACAAAAAAGAAA 25610
 QY 59 ACTATAGACCAAGTACCACTGATGTAATTAATCAAGAAATCCCAACAAATCTAGCTA 118
 Db 25611 ACTACAGATCAATATCTTGTAAACATAGATGTAATTAATCTTAACAAATCTAGCTG 25670
 QY 119 ACCCAATCCAACAGATATCAAGAAATTAATCAACATGTCAGAGTGGTTCATACAG 178
 Db 25671 ACTGAATCCAACAGATATCAAGAAATTAATCAACATGTCAGAGTGGTTCATACAG 25730
 QY 179 GGGTGCAG-GATAGGTTAATCAATCAACAGATCAATTAATGTAATCATCAATTAACAGAA 237
 Db 25731 GGGTGCAGAGATGTTAATGTAATCAACAGATCAATTAATGTAATCATCAATTAACAGAA 25790
 QY 238 TTAATAAACAATAACATATATCATCTCAATAGATGCTGAAAAAGCACTTGAACAAATCT 297
 Db 25791 TTAATAAACAATAACATATATCATCTCAATAGATGCTGAAAAAGCACTTGAACAAATCT 25850
 QY 298 AACATTTCTTATGATTAATTAATCACTTCAACAAATCAATGAAGACATACCTTAATG 357
 Db 25851 AGCATCCCTTATATTAATTAATCACTTCAACAAATCAATGAAGACATACCTTAATG 25910
 QY 358 TAATTAAGCCATATATGACGGACCAACAGCAACATTAATCTGAATGGGAAAAAGTTGA 417
 Db 25911 TAATTAAGCCATATATGACGGACCAACAGCAACATTAATCTGAATGGGAAAAAGTTGA 25970
 QY 418 AACATTTCTTATGATTAATTAATCACTTCAACAAATCAATGAAGACATACCTTAATG 476
 Db 25971 AAGATTTCTTATGATTAATTAATCACTTCAACAAATCAATGAAGACATACCTTAATG 26030
 QY 477 ACATAGTAGTGAAGTTTGAAGCCAGCAATCAAGAAAGAAATCAAGGCAACCC 536
 Db 26031 ACATAGTAGTGAAGTTTGAAGCCAGCAATCAAGAAAGAAATCAAGGCAACCC 26090
 QY 537 AAATCAATTAAGAGAGAGTCAATCTGCTTCACTGATGATATGATATACCTAG 596
 Db 26091 AAATCGTAAAGAGAGAGTCAATCTGCTTCACTGATGATATGATATACCTAG 26150
 QY 597 AAAACCTTAAGAGTCAATCTGCTTCACTGATGATATGATATACCTAG 656
 Db 26151 AAAACCTTAAGAGTCAATCTGCTTCACTGATGATATGATATACCTAG 26210
 QY 657 CAGGATCAAACTAAATGTAACAACAAATCAGTAGCACTATACCAACAGTACCAAG 716
 Db 26211 CTGATTCAAAGATTAATGTAACAACAAATCAGTAGCTCTTCTATATACCAACAGTACCAAG 26270
 QY 717 CTGAGATCAAACTAAAGAGTCAATCTTCAATATGCTGTAAA-----AAA 765

/product= "Human MrgX3 protein"

XX XX MO200183555-A2.

XX PD 08-NOV-2001.

XX PF 04-MAY-2001; 2001MO-US014519.

XX PR 04-MAY-2000; 2000US-0202027P.

XX PR 01-AUG-2000; 2000US-0222344P.

XX PR 03-NOV-2000; 2000US-00704707.

XX PR 19-APR-2001; 2001US-0285493P.

XX PA (CALY) CALIFORNIA INST OF TECHNOLOGY.

XX PI Anderson DJ, Dong X, Zylka M, Han S, Simon M;

XX DR WPI; 2002-171346/22.

XX DR P-PSDB; MAF21296.

XX PT Isolated polypeptide, Mrg, which is a G-protein coupled receptor and an

XX PT isolated polypeptide, dtr-12, which is also a receptor, useful for

XX PT identifying agonists or antagonists for treating pain.

XX PS Disclosure; Page 128-129; 185pp; English.

XX CC The invention relates to Mrg (mas-related gene) protein, which is a G-

XX CC protein coupled receptor and dtr-12 protein, which is a receptor. The

XX CC invention is useful for identifying compounds that bind to it, especially

XX CC agonists or antagonists. Administration of an agent (e.g. the identified

XX CC agent) that increases the expression of Mrg in a mammal may be used for

XX CC treating impaired sensory perception in a mammal, especially pain. The

XX CC antagonist may also be useful for treating impaired sensory perception in

XX CC a mammal. The present sequence is human MrgX3 DNA

XX SQ Sequence 1400 BP; 276 A; 375 C; 363 G; 386 T; 0 U; 0 Other;

Query Match 15.9%; Score 1374.6; DB 6; Length 1400;

Best Local Similarity 99.3%; Pred. No. 9.1e-310;

Matches 1391; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 7196 TCAGGCCAGGATGAGTATCATCGGGTCACAGCACTGCTGATGAGTGGGGTGT 7255

DB 1 TCAGGCCAGGATGAGTATCATCGGGTCACAGCACTGCTGATGAGTGGGGTGT 60

QY 7256 TTGATCTTAATGTATCCCATGTACAGCAGAACTTGTGGCAGTAGAGAGTCTAG 7315

DB 61 TTGATCTTAATGTATCCCATGTAGCAGAACTTGTGGCAGTAGAGAGTCTAG 120

QY 7316 GCTTCAGAGTCAACAAGAACTGATTTCAAACTGAGATTGAGACCCCACTTTTGATA 7375

DB 121 GCTTCAGAGTCAACAAGAACTGATTTCAAACTGAGATTGAGACCCCACTTTTGATA 180

QY 7376 GGATCTTAATGTATCCCATGTACAGTCTCTCTCTTTAATAGAGCACTAATATCCCA 7435

DB 181 GGATCTTAATGTATCCCATGTACAGTCTCTCTCTTTAATAGAGCACTAATATCCCA 239

QY 7436 CATGGCAGGAGTGTGGGAGATCAGAGATCAAAAGCTGTGATCATCTGTTCTG 7495

DB 240 CATGGCAGGAGTGTGGGAGATCAGAGATCAAAAGCTGTGATCATCTGTTCTG 299

QY 7496 TTTTCAGAGGTATCAAGCTGGGGTTTCTGAGCATGAGATTCAACCAATCCCACTTTGGGTA 7555

DB 300 TTTTCAGAGGTATCAAGCTGGGGTTTCTGAGCATGAGATTCAACCAATCCCACTTTGGGTA 359

QY 7556 CAGAACTGACACCAATCAAGGAGAGCTGCTGTCTACAGAGCACTCTGAGCT 7615

DB 360 CAGAACTGACACCAATCAAGGAGAGCTGCTGTCTACAGAGCACTCTGAGCT 419

QY 7616 TCACGGGGCTGACGTCGATCTTTCCCTTGTGCGCTGACAGAAACCGGTTGTGCTCT 7675

DB 420 TCACGGGGCTGACGTCGATCTTTCCCTTGTGCGCTGACAGAAACCGGTTGTGCTCT 479

QY 7676 GGCTCTGGGCTGCGGAGATGCGGAGAAAGCTGTCTCCATCTACATCTCTCAACCTGTGCG 7735

DB 480 GGCTCTGGGCTGCGGAGATGCGGAGAAAGCTGTCTCCATCTACATCTCTCAACCTGTGCG 539

QY 7736 CGGCGGACTTCTCTCTCTTGTAGCGGCACTTATATGTGTGCGGTTAGCCCTCATCATTA 7795

DB 540 CGGCGGACTTCTCTCTCTTGTAGCGGCACTTATATGTGTGCGGTTAGCCCTCATCATTA 599

QY 7796 TCCGCAATCCCATCTCCAAAATCTCAAGTCTGTGATGACCTTTTCCCTTATATAGGCC 7855

DB 600 TCCGCAATCCCATCTCCAAAATCTCAAGTCTGTGATGACCTTTTCCCTTATATAGGCC 659

QY 7856 TAAAGATGCTAGAGGCGCATCAGCACCGAGCCCTGCTGTCCATCTGTGCGCCATCTGCT 7915

DB 660 TAAAGATGCTAGAGGCGCATCAGCACCGAGCCCTGCTGTCCATCTGTGCGCCATCTGCT 719

QY 7916 ACGACTGCGCGCGCGCGCGAGATACGTCATGAGTATGTCGTCGTCGTCGCGCGCGCTGT 7975

DB 720 ACGACTGCGCGCGCGCGCGAGATACGTCATGAGTATGTCGTCGTCGTCGTCGCGCGCTGT 779

QY 7976 CCTGCTGCGGAGTATCTGAGAGTGTCTGTGACTTCTCTGTTTATGAGTGTGAT 8035

DB 780 CCTGCTGCGGAGTATCTGAGAGTGTCTGTGACTTCTCTGTTTATGAGTGTGAT 839

QY 8036 CTGTTTGTGTGAAAACGTCAATTTTCAATACATCGCTGCTGCTTTTATATGTGCG 8095

DB 840 CTGTTTGTGTGAAAACGTCAATTTTCAATACATCGCTGCTGCTTTTATATGTGCG 899

QY 8096 TTTCTGTGAGGTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8155

DB 900 TTTCTGTGAGGTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 959

QY 8156 CGCTGACCAAGCTGTAGCACTGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8215

DB 960 CGCTGACCAAGCTGTAGCACTGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1019

QY 8216 TGCCCTTTGGGATTCAGTGGGCTCTGTTTTCAGAGATCACCTGATTTGAAAGCTTAT 8275

DB 1020 TGCCCTTTGGGATTCAGTGGGCTCTGTTTTCAGAGATCACCTGATTTGAAAGCTTAT 1079

QY 8276 TTTGTCAATGACATCTGATTTTCAATTTTCTGTCGCTTCAACAGAGTGCACCCCA 8335

DB 1080 TTTGTCAATGACATCTGATTTTCAATTTTCTGTCGCTTCAACAGAGTGCACCCCA 1139

QY 8336 TCATTTACTTCTCTGTGGGCTCTTTTAGGACGCTCAAAATAGGACAGAACTGAGCTG 8395

DB 1140 TCATTTACTTCTCTGTGGGCTCTTTTAGGACGCTCAAAATAGGACAGAACTGAGAGCTG 1199

QY 8396 TTTTCAGAGGCTCTGACAGACACGCTTGAAGTGTGAAAGTGTGAGGCTGCTCTC 8455

DB 1200 TTTTCAGAGGCTCTGACAGACACGCTTGAAGTGTGAAAGTGTGAGGCTGCTCTC 1259

QY 8456 AGGAAACCTTGAGCTGTGCGGAGAGAGATTTGAGAGTGTGAGAGAACTCTGCGCTGT 8515

DB 1260 AGGAAACCTTGAGCTGTGCGGAGAGAGATTTGAGAGTGTGAGAGAACTCTGCGCTGT 1319

QY 8516 CAGACAGACTTGTAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8575

DB 1320 CAGACAGACTTGTAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1379

QY 8576 TAGCCTTCTGCTCAAAATG 8596

DB 1380 TAGCCTTCTGCTCAAAATG 1400

RESULT 10

ABZ42595

ID ABZ42595 standard; DNA; 1400 BP.

XX AC ABZ42595;

XX DT 04-MAR-2003 (first entry)

DE Human G protein-coupled receptor MrgX3 nucleotide seq ID NO:673.
 XX G protein-coupled receptor, GPCR; antigenic peptide; gene therapy;
 KM G protein-coupled receptor modulator; antibody; immune-related disease;
 KM growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KM immunological-related disease; cell proliferative disease; autoimmune disease;
 KM Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KM osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
 KM graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KM psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 KM mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KM hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KM ulcer; gene; ds.
 XX Homo sapiens.
 OS WO200261087-A2.
 PN 08-AUG-2002.
 PD 19-DEC-2001; 2001WO-US050107.
 PF 19-DEC-2001; 2000US-0257144P.
 PR 19-DEC-2000; 2000US-0257144P.
 XX (LIFE-) LIFESPAN BIOSCIENCES INC.
 PA Burner GC, Roush CL, Brown JP;
 FI WPI; 2003-046718/04.
 DR P-PSDB; ABB81750.
 XX New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR), useful for diagnosing and designing drugs for treating conditions
 PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
 PT autoimmune diseases.
 PS Disclosure; Fig 1; 523pp; English.
 XX The present invention describes antigenic peptides (I) comprising: (a)
 CC any one of 1601 sequences (see ABB82019 to ABB83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular G
 CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity or
 CC avidity for a particular GPCR. (I) can be used as GPCR modulators and in
 CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
 CC antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCRs. The antigenic peptides for
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related disease, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcer, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnosis. ABB82523 to ABB82869 encode
 CC GPCR proteins given in ABB81675 to ABB82018, which are used in the
 CC exemplification of the present invention
 XX
 XX Sequence 1400 BP; 276 A; 375 C; 363 G; 386 T; 0 U; 0 Other;
 SO
 Query Match 15.9%; Score 1374.6; DB 8; Length 1400;
 Best Local Similarity 99.3%; Pred. No. 9.1e-310;
 Matches 1391; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
 QY 7186 TCAGGCCAGATAGATATATCATCGGCTCCACAGCACTGCTAGATGAGTGCGGCTGT 7255
 DB 1 TCAGGCCAGATAGATATATCATCGGCTCCACAGCACTGCTAGATGAGTGCGGCTGT 60

QY 7256 TTGATCTTAATGTTATCCCATGTGACAGACAGAACTGTGTGAGAGAGAGTCCG 7315
 DB 61 TTGATCTTAATGTTATCCCATGTGACAGACAGAACTGTGTGAGAGAGAGTCCG 120
 QY 7316 GCTTCAGAGTCAACAGAACTGGAATTCAACTGGAATTTAGAGACCCCACTTTGATA 7375
 DB 121 GCTTCAGAGTCAACAGAACTGGAATTCAACTGGAATTTAGAGACCCCACTTTGATA 180
 QY 7376 GGTGACTTATTTCTGTGAGAGTCTGTATCTCTCTTTAAATGAGAGACGTAATCCCA 7435
 DB 181 GGTGACTTATTTCTGTGAGAGTCTGTATCTGTGAGAGTCTCTTTAAATGAGAG-AGTAATATCCCA 239
 QY 7436 CATGGAGAGGTGTGGGGAGATCAGAGATCAAAAGCTGTGATCACTGTGTTCTG 7495
 DB 240 CATGGAGAGGTGTGGGGAGATCAGAGATCAAAAGCTGTGATCACTGTGTTCTG 299
 QY 7496 TTTCAGAGGTCAATCAGACTGAGGTTTCTAGAGATGGAATCAACATCCAGTCTTGAGTA 7555
 DB 300 TTTCAGAGGTCAACAGACTGAGGTTTCTAGAGATGGAATCAACATCCAGTCTTGAGTA 359
 QY 7556 CAGACTGACCAATCAACAGAGAGTGTGAGAGACTCTTGTCTCAAGACGACCTGTAGCT 7615
 DB 360 CAGACTGACCAATCAACAGAGAGTGTGAGAGACTCTTGTCTCAAGACGACCTGTAGCT 419
 QY 7616 TCAGGGGGGTGAGGTGATCGTTTCCCTGTGCGGTGACAGGAAACGGGTTGTGCTCT 7675
 DB 420 TCAGGGGGGTGAGGTGATCGTTTCCCTGTGCGGTGACAGGAAACGGGTTGTGCTCT 479
 QY 7676 GGTCTCTGAGGTGAGGTGATCGGTGAGAGAGAGTGTCTCATCTCAATCTCTCAACCTGTGCTG 7735
 DB 480 GGTCTCTGAGGTGAGGTGATCGGTGAGAGAGAGTGTCTCATCTCAATCTCTCAACCTGTGCTG 539
 QY 7736 CGGCGCACTTCTCTCTCTTCTTGTAGGAGCAATTAATGAGTGTGAGTGTGAGTGTGAGT 7795
 DB 540 CGGCGCACTTCTCTCTCTTCTTGTAGGAGCAATTAATGAGTGTGAGTGTGAGTGTGAGT 599
 QY 7796 TCCGCAATCCCATCTCAAAATCTCAGTCTGTGATGACCTTCCCTTCTTATATAGGC 7855
 DB 600 TCCGCAATCCCATCTCAAAATCTCAGTCTGTGATGACCTTCCCTTCTTATATAGGC 659
 QY 7856 TAAGCATGTGAGAGGATCAGACAGAGAGTGTGATGATCTGTGAGTGTGAGTGTGAGT 7915
 DB 660 TAAGCATGTGAGAGGATCAGACAGAGAGTGTGATGATCTGTGAGTGTGAGTGTGAGT 719
 QY 7916 ACACATGCG 7975
 DB 720 ACACATGCG 779
 QY 7976 CCGTGTGCGGAGATATCTGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 8035
 DB 780 CCGTGTGCGGAGATATCTGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 839
 QY 8036 CTGTGTGTGTGAAGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 8095
 DB 840 CTGTGTGTGTGAAGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGAT 899
 QY 8096 TTCTGT 8155
 DB 900 TTCTGT 959
 QY 8156 CGCTGACAGGCTGT 8215
 DB 960 CGCTGACAGGCTGT 1019
 QY 8216 TGCCCTTTGTGATTCAGTGTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8275
 DB 1020 TGCCCTTTGTGATTCAGTGTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1079
 QY 8276 TTGTGATGTGATCTAGT 8335
 DB 1080 TTGTGATGTGATCTAGT 1139

QY 8336 TCATTATCTTCTGTCGTCCTTTAGGACGCTCAAAATAGGACGAACTGTAAGCTGG 8395
 DB 1140 TCATTATCTTCTGTCGTCCTTTAGGACGCTCAAAATAGGACGAACTGTAAGCTGG 1199
 QY 8396 TTCTCCAGAGGGCTCTGACAGACGCTGAGTGAAGTGAAGGCTGCTTCTC 8455
 DB 1200 TTCTCCAGAGGGCTCTGACAGACGCTGAGTGAAGTGAAGGCTGCTTCTC 1259
 QY 8456 AGGAAACCTTGAGCTGTGAGGAGCAGATTGAGCAGTGAAGAACTCTGCTGT 8515
 DB 1260 AGGAAACCTTGAGCTGTGAGGAGCAGATTGAGCAGTGAAGAACTCTGCTGT 1319
 QY 8516 CAGACAGACTTGAAGCAATGCTGCTGCACTTGAATTAATGATTTTCT 8575
 DB 1320 CAGACAGACTTGAAGCAATGCTGCTGCACTTGAATTAATGATTTTCT 1379
 QY 8576 TAGCTTCTGCTCAGAAATG 8596
 DB 1380 TAGCTTCTGCTCAGAAATG 1400

RESULT 11
 ADH08534
 ID ADH08534 standard; DNA; 1400 BP.
 AC ADH08534;
 DT 25-MAR-2004 (first entry)
 DE DNA sequence MrgX3.
 KW mas-related gene D; MrgD; Analgesic; Vulnerary; Ophthalmological;
 KM sensory perception; glaucoma; Mrg; ds.
 OS Mus musculus.
 PN W02004003133-A1.
 PD 08-JAN-2004.
 PF 13-MAY-2003; 2003MO-US015004.
 PR 26-JUN-2002; 2002US-00183116.
 PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
 PI Anderson DJ, Dong X, Zylka M, Han S, Simon MI;
 DR WPI; 2004-083025/08.
 DR P-PSDB; ADH08535.

New mas-related gene D polypeptides, useful as therapeutics or in identifying agonists or antagonists that alter pain perception in a mammal for treating impaired sensory perception, e.g. chronic intractable pain or neuropathic pain.

The present invention relates to an isolated mas-related gene D (MrgD) polypeptide. The MrgD polypeptides are useful as therapeutics or for identifying compounds, i.e. agonists or antagonists, that alter pain perception in a mammal. The compounds are useful for treating impaired sensory perception, e.g. chronic intractable pain or neuropathic pain, promoting wound healing, restoring normal sensitivity following injury, or treating ocular conditions, particularly those associated with pressure such as glaucoma. The Mrg genes or proteins may be used as molecular probes for the detection of cells or tissues related to or involved with sensory perception. The present sequence represents a MrgA (Mrg subfamily) encoding sequence.

Sequence 1400 BP; 276 A; 375 C; 363 G; 386 T; 0 U; 0 Other;
 Query Match 15.9%; Score 1374.6; DB 12; Length 1400;

Best Local Similarity 99.3%; Pred. No. 9.1e-310;
 Matches 1391; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 7196 TCAGGCCAGAGTGAAGTAATCATCGGGTCCACAGACATGCTAGATGAGTGGGGTGT 7255
 DB 1 TCAGGCCAGAGTGAAGTAATCATCGGGTCCACAGACATGCTAGATGAGTGGGGTGT 60
 QY 7256 TTGATCCTAATGTTATCCCATGTCAGACAGAACTGTGTGGCAGTGAAGAGTCAAG 7315
 DB 61 TTGATCCTAATGTTATCCCATGTTAGACAGAACTGTGTGGCAGTGAAGAGTCAAG 120
 QY 7316 GCTTCAGAGTCAAGAAAGCTGAAATTTCAACTGGAATTTGAGACCCCACTTTGATA 7375
 DB 121 GCTTCAGAGTCAAGAAAGCTGAAATTTCAACTGGAATTTGAGACCCCACTTTGATA 180
 QY 7376 GGTGACTTATTTCTGTCAGAGTCTGTCATCTCTCTTTAATGAGACAGTAAATCCCA 7435
 DB 181 GGTGACTTATTTCTGTCAGAGTCTGTCATCTCTCTTTAATGAGACAGTAAATCCCA 239
 QY 7436 CATGGCAGGGTGTGGGGAGAAATGAGATCAAAAGCTGGTATCATATCTGTTCTG 7495
 DB 240 CATGGCAGGGTGTGGGGAGAAATGAGATCAAAAGCTGGTATCATATCTGTTCTG 299
 QY 7496 TTTCAGGGTATCAGACTGGGGTTTCTGACATGAAATTCACCATCCAGTCTTGGATA 7555
 DB 300 TTTCAGGGTATCAGACTGGGGTTTCTGACATGAAATTCACCATCCAGTCTTGGATA 359
 QY 7556 CAGAACTGACACCAATCAACGAGATGAGAGACTCTCTTGTCAAGACAGACCTGAGCT 7615
 DB 360 CAGAACTGACACCAATCAACGAGATGAGAGACTCTCTTGTCAAGACAGACCTGAGCT 419
 QY 7616 TCAGGGGCTGACGTGATCGTTTCCCTGTGCGCTGACAGAAACGCGGTGTGCTCT 7675
 DB 420 TCAGGGGCTGACGTGATCGTTTCCCTGTGCGCTGACAGAAACGCGGTGTGCTCT 479
 QY 7676 GGTCTCTGGGGTGGCGGATGCGGAGAAAGCTGTCTCATCTTAATCTCAACCTGTGCTG 7735
 DB 480 GGTCTCTGGGGTGGCGGATGCGGAGAAAGCTGTCTCATCTTAATCTCAACCTGTGCTG 539
 QY 7736 CGGCCGACTTCTCTTCTTGTAGCGGCACATTAATGTTGCGCTTACGCTCATCAATA 7795
 DB 540 CGGCCGACTTCTCTTCTTGTAGCGGCACATTAATGTTGCGCTTACGCTCATCAATA 599
 QY 7796 TCGGCATCCCATCTCCAAAATCTCACTGTGATGACCTTTCCCTAATTAAGGCC 7855
 DB 600 TCGGCATCCCATCTCCAAAATCTCACTGTGATGACCTTTCCCTAATTAAGGCC 659
 QY 7856 TAAACATCTGAGGGCCATGACACGAGCGCTGCTCATCTGTGGCCCATCTGGT 7915
 DB 660 TAAACATCTGAGGGCCATGACACGAGCGCTGCTCATCTGTGGCCCATCTGGT 719
 QY 7916 ACCACTCCCGCCGACAGATACCTGTGATGAGTGAATGAGTGTGCTGTGAGGCCCTGT 7975
 DB 720 ACCACTCCCGCCGACAGATACCTGTGATGAGTGAATGAGTGTGCTGTGAGGCCCTGT 779
 QY 7976 CCTGTGCGAGATATCTGAGTGAATGTTCTGTGACCTTCTGTTAGTGTGCTGAT 8035
 DB 780 CCTGTGCGAGATATCTGAGTGAATGTTCTGTGACCTTCTGTTAGTGTGCTGAT 839
 QY 8036 CTGTTTGTGTGAAGACATCAATTTCAATTAACCGGTGCTGCTTTTATGCTGG 8095
 DB 840 CTGTTTGTGTGAAGACATCAATTTCAATTAACCGGTGCTGCTTTTATGCTGG 899
 QY 8096 TTCTCTGTGGTCAAGCTGTGCTGTCGTGACAGATTTCTGTGATCCCGAAGATGC 8155
 DB 900 TTCTCTGTGGTCAAGCTGTGCTGTCGTGACAGATTTCTGTGATCCCGAAGATGC 959
 QY 8156 CGCTGACAGGCTGATGATCAATCTCTGACAGTGTGCTTCTCTCTGTGCTG 8215
 DB 960 CGCTGACAGGCTGATGATCAATCTCTGACAGTGTGCTTCTCTCTGTGCTG 1019
 QY 8216 TGCCCTTTGATTCATGTGGGCTGTGTTTCCAGATTCACCTGATTTGAATACTTAT 8275


```

Qy 1361 AGATGATCAAGACCTTAATCTGAGCCTTAAACATTAATAATTTCTAGAGATTAACATC 1420
Db 26913 AGATGATTAAGGCTTAATCAATTAAGAGTAACTAT-GAAATTTTAAAGATTAACCTT 26971
Qy 1421 AGAAAAATGCTTACAGACATTCATAGGCAAAAGATTCATGCGCAAGAACCCAAAAGTA 1480
Db 26972 GGAABAAACCTTCTAGACATTTGGCTTAGGCAAGATTTCTATGACCAAGAACCCAAAAGCA 27031
Qy 1481 AATGCAACAAAAAATAATTAATAGATAGACTTAATTAAC-TAAAAAGCTTTTGGGC 1539
Db 27032 AATGCAATAAAAAATAATTAATAGCTGGGACCTCATTAACCTTACAGACCTTTTGAG 27091
Qy 1540 AGCAAAAACATCTATTGACAGCAACACCAACCCAGTGAAG-AAAATCTTCA 1598
Db 27092 GGGCAAAAGAACACTGACAGAGTAAACAGCAACCAACAGTGGGAGAAAAATCTTCA 27151
Qy 1599 CAAACTAGCATCTGACTAAGGACTAATATCCGGAATCCAGAAAGAACTCAAAATCA 1658
Db 27152 CAATCTATCTCTGACAAAGG-GTAGATTCAGAAATCTCAAGAGACCCCAAAATCA 27210
Qy 1659 GCAAGAAAGAAAGCAAAACATCCATGAAGAGTGGGCTAAGGACATGATAGCAATTC 1718
Db 27211 GTAAAGAAAAAACAACATCCATCAAAAAGTAGGGCTAAGGACATGATAGCAATTC 27270
Qy 1719 CAAAAGAGATATCAATAGGCCCAACAC--AGAAAAATGCTTAACATCACTAATGA 1776
Db 27271 CAAAAGAGATATCAATAGGCCCAACACATAGAAAAAAATGCTCAACATCACTAATGA 27330
Qy 1777 TTAGGGAATATGTAATCAACA-CTGTAATGCGATACACCTTCTCCGTCGAAGATGGT 1834
Db 27331 TCAGAGAAATGCAATCAAAAACAAATATGATACACCTTCTCTGCAAGATGGC 27390
Qy 1835 CATATTTTAAAAATCT--AAAAATATAGATGTTGGTGGTCTGTGTGATTAAGAA 1892
Db 27391 CATATTTTAAAAATCTTAAAAAAGTAGATGTTGGATGGAAGGGGGATCAAGAAACA 27450
Qy 1893 CTTTAACTGCTGCTGGGAATGTAACCTTGGCAACATATAGAAAAACAGTGTGAAA 1952
Db 27451 CTTTAACTGCTGCTGGGAATGTAACCTTGGCAACATATAGAAAAACAGTGTGAAA 27510
Qy 1953 TTTCTTAAGGAATCAAAAGTATGATGACCATTTGATCAGCAATCCCAT----- 2001
Db 27511 TTTCTTAAGGAATCAAAAGTATGATGACCATTTGATCAGCAATCCCAT----- 27570
Qy 2002 ----- 2001
Db 27571 TACCCAGAGAAAGATGATTTATTTGAAAAAGACCTTGACAGATTTATAGCA 27630
Qy 2002 ----- 2001
Db 27631 GCACAAATTCACAAATTCGAAAACTGTGAACCTAACCAATGTCATCAGTCAATTAAGTG 27690
Qy 2002 -TAAATATGTAATAATATATATATTTATATTCATGGAATCAACTCAGCATATAAAG 2060
Db 27691 AATAAGAACTGTGGTACAGATATATATCAATGAAATCTACACAGCTATGAAAAAG 27750
Qy 2061 AATAAAATGATGATTCACAGCAATCTAGATGGAATGAGACCTTAATCTAAGTGG 2120
Db 27751 AATGAATTAACAGATTTGACAGTGAAGTGAATTTGAGACATTTATTTCTAAGTGA 27810
Qy 2121 GTAACTCAGGATGAAAAACCAACATCATATGTTCTCACTTACAAAGTGGGGCTTAAGT 2180
Db 27811 GTAACTCAGGATGAAAAACCAACATCATATGTTCTCACTGATATGAGATTAAGT 27870
Qy 2181 GTGAGACAGAAAGCAT-AGAATGATATTAATGAATCTGAGGACTTGAAGGGGAAAGATG 2239
Db 27871 ATGAGACACAAAGATTAAGAAATGATTAATGAATCTTGGAGACTTGGGGGAAAGATG 27930
Qy 2240 GAAAGAGGCGAGGATTAAGAAATGATTAATGAATCTTGGAGACTTGGGGGAAAGATG 2299
Db 27931 GAGAGGGGGTGAAGGATTAAGAAATGATTAATGAATCTTGGAGACTTGGGGGATG 27990
Qy 2300 GGAGACCAAAATCTCAGAAATTAACCACTTAAGAACTTATCATGAGAAACA--ACACA 2357

```

```

Db 27991 GGTGACCAAAATTCACAAATCCACCTAAGAACTTACTATATATCAATTAATACCA 28050
Qy 2358 CCGTGTCCCAAAATCCCAATGAATAATTAATTAATTAATTAATTAATTAATTCACAG 2417
Db 28051 CCGTGTCCCTAATTAATTAATGAABAAAAATTTTATTAAGTAATTAATTAATTAATTA 28110
Qy 2418 AATT 2421
Db 28111 GTTT 28114

RESULT 13
ABN96931/c
ID ABN96931 standard; DNA; 99014 BP.
XX
AC ABN96931;
XX
DT 13-AUG-2002 (first entry)
XX
DE Gene #3429 used to diagnose liver cancer.
XX
KW Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS Homo sapiens.
XX
PN MO200229103-A2.
XX
PD 11-APR-2002.
XX
PE 02-OCT-2001; 2001WO-US030589.
XX
PR 02-OCT-2000; 2000US-0237054P.
XX
PA (GENE-) GENE LOGIC INC.
XX
PI Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX
DR WPI; 2002-426119/45.
XX
PT Diagnosing and detecting the progression of liver cancer, hepatocellular
PT carcinoma or metastatic liver tumor in a patient, involves detecting the
PT level of expression of two or more genes in a liver tissue sample.
XX
PS Claim 1; SEQ ID NO 3429; 298bp; English.
XX
CC The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytosolic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 99014 BP; 27179 A; 21575 C; 21336 G; 28924 T; 0 U; 0 Other;
XX
Query Match 15.8%; Score 1358; DB 6; Length 99014;
Best Local Similarity 78.6%; Pred. No. 5.6e-305;
Matches 1895; Conservative 0; Mismatches 345; Indels 170; Gaps 16;
Qy 131 AGCATATCAAGAAATATATCAATTCATGTAAGTGGGTTTCATACAGGGGTGAGGA-T 189
Db 88976 AGAAAAAATTAAGATATATCAATTCATGTAAGTGGGTTTCATACAGGGGTGAGGAGT 88917

```

QY 190 AGTTAATACATACAGTGAATTAATGTGATACATGACATTAACAGAAATTAATAACAAA 249
 Db 88916 GGTTTTAAACATATGCAAGTTAAATGATGATATACACATTAACAGAAATTAATAACAAA 88857
 QY 250 ATCAATGATCATCTCAATAGATGTGTAATAAGCAATTTGACAAATCTTAATCTTTA 309
 Db 88856 ATCAATGATCATCTCAATAGATGTGTAATAAGCAATTTGACAAATCTTAATCTTTA 88797
 QY 310 TGATTTAAACCTTCAGCAAAATGACATAGAAAGACATACCTTAATGTATTAAGGCA 369
 Db 88796 TGATTTAAACCTTCAGCAAAATGCGCATACAAAGCATCTCAATGTATTAAGGCA 88737
 QY 370 TATATAGCGAGCCACAGCAAAATATATCTGTAATGGGAAAAGTTGAAAACATTTGCC 429
 Db 88736 TCTATAGCGAACCCACAGCCAAATATATCTGTAATGGGAAAAGTTGAAAACATTTGCC 88677
 QY 430 TGAGAACTGGAACAGACAAAGATG-CTACTTTCACTTCTTATTCACATAGTAGTG 488
 Db 88676 TGAGAACTGGAACAGACAAAGATGCCACTCTCACTACTCTTCAATAGTAGTG 88617
 QY 489 AAGTTTACGAGAGCAATCAGACAGAGAAAGAAATCAAGGCAACCAATCAATTAAG 548
 Db 88616 AAGTTTACGAGAGCAATCAGACAGAGAAAGAAATCAAGGCAATCC-AATTGTAAG 88558
 QY 549 AGGAATGCAACTGTCTCTTCTTCACTGATGATATGATGTATACCTAGAAACCTTAAG 608
 Db 88557 AGGAATGCAACTGTCTCTTCTTCTGATGATATGATGTATACCTAGAAATCTTAAG 88498
 QY 609 ACTCATCCAGAAAGCTCTAGAACTGATACATTAATTCAGTAAGTTTCAGATATCAAC 668
 Db 88497 ACTCATCCAGAAAGCTCTAGAACTGATATTAATAAGTAAGTAAGTTTCAGATATCAAC 88438
 QY 669 TAAATGTACAAATAGTAGAGCTGTATACCAACAGTGAACCAAGTGAAGATCAAA 728
 Db 88437 TAAATGTACAAATAGTAGAGCTCTTATATACCAACAGTGAAGATCAAA 88378
 QY 729 TCAAGAACCTCAACACTTTTACATAGCTGTAAA-----AAATCTTGAAGATATCT 782
 Db 88377 TCAAGAACCTCAACCTTTTACATAGCTGTCAAAAAGATTAATTAATCTTGAAGATATCT 88318
 QY 783 TACCCAGAGAGGTGAAGACCTCTCAAGAGAAATCTCAAAACAGACT----GACATCA 838
 Db 88317 TACCCAGAGAGGTGAAGACCTCTCAAGAGAAATCTCAAAACAGCTGTGAAGAAAGCA 88258
 QY 839 TAGATGACAAACAAAGTGAACACATCCCATGCTCATGATGGGTAGATCAATATG 898
 Db 88257 TAGATGACAAACAAAGTGAACACATCCCATGCTCAAGATGGGTAGATCAATATG 88198
 QY 899 TGAATAATGCAATATGCGCAAAAGCAATCTAAGATGCAATGCAATCCCAACAAATAT 958
 Db 88197 TGAATAATGCAATCTTCCAAAGCAATCTAAGATGCAATGCAATCCCAACAAATAT 88138
 QY 959 CATCATCATCTTCACAGAACTAGAAAAAAACAATTTCAAAATTCATATGAAACAAAC 1018
 Db 88137 CATCATCATCTTCACAGAACTAG-AAAAACAATTTCAAAATTCATATGAAAC----- 88085
 QY 1019 CAAAAAATGCAATGCGCAATGCGCAAGCAAGCACTTAGCAAAAAGCAAAATCTG 1078
 Db 88084 -----CAAAAAAGGCCCAATGCGCAAGCAAGCAATGCAAAAAGCAAAATCTG 88033
 QY 1079 GAGGATGCACTTACCATCTTCAAACTATATCAAGGCTATTAATCAACAAACATCAT 1138
 Db 88032 GAGGATGCACTTACCATCTTCAAGCTATATTAAGGCTATTAAGCAAAACAGCAT 87973
 QY 1139 GGCATGACATTAATAGGCAATAGCAATGAAAGAAAGAGAAATCCAGAAATTA 1198
 Db 87972 GGCATGACATTAATTAAGCAATAGCAATAGCAATGAAAGAAAGAGAAATCCAGAAATTA 87913
 QY 1199 AGCCAAATTAATTAAGCAATAGATTTTGAACAAAGCAACAAACATTAAGTGGGAA 1258
 Db 87912 ACCCAATTAATTAATTAAGCTATCTTCAACAAAGCAACAAACATTAAGTGGGAA 87853

QY 1259 AAGACATCTAGTAAACAATGATGCTGAGATTAATGGAAGCCACATGTGGAAGATGA 1318
 Db 87852 AAGACATCTTTCACAAATATGATGCTGGATTAATGCTGACCAATGTGCAAGATGA 87793
 QY 1319 AACTGATCCCTGTCTCTCACTTAATACAAAATTTGATTAAGATGATCAAGACTTA 1378
 Db 87792 AACTGAAT-CTTCATCTCTCACTTAATACAAAATTTCAAACTCAAGATGATTAAGACTTA 87734
 QY 1379 AATCTGAGACCTTAACCAATTAATAATCTAGAAATTAATCAAGAAATGCTGTAGAC 1438
 Db 87733 AATCTGAGACCTTAACCAATTAATAATCTAGAAATTAATCAAGAAATGCTGTAGAC 87674
 QY 1439 ATTCACTAGGCAAGACTTCAATGCGCAAGAACCCAAAGTAATGCAACAAACAAA 1498
 Db 87673 ATTGCTTAGGCAAGACTTCAATGCAAGAACCCAAAGTAATGCAATTAAGCAATTAAG 87614
 QY 1499 ATAAATAGATAGGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1558
 Db 87613 ATAAATAGCTGGAC-----TTAAACTAAGACTTTTGCACGAAAAAGCAAGTCAAGC 87558
 QY 1559 AGAGCAACAGACACACCCACCGATGAGGAAATTTCAACAACTAAGCATGACTTA 1618
 Db 87557 AGGTAA--AGAGAACCCACAGATGGGAGAAAGCTTCACAATTAATCACTTGAACAA 87500
 QY 1619 GCACTAATATCCGGAATCCACAAAGAACTCAACAAATGACGAGAGAAAGCAAAAT 1678
 Db 87499 GCAATTAATCCGGAATCTCAATGAACTCAACAAATGACGAGATTAAGAAACAAAT 87440
 QY 1679 CCCATGAAGAGTGGCTAAGACATGAATAGCAATTTCTCAAAAGAGATTAACAAATG 1738
 Db 87439 TCCATGAAGAGTGGCTAAGACATGAAGAACTCAAAAGAGATTAACAAATG 87380
 QY 1739 GCCAACAACAGAAAAAATGCTTAACATCACTAATGATTAAGGAAATGTAATCAAC 1798
 Db 87379 -----CATATGAATAAATGCTCAACATCAACCAATGCGGAAATGCAATCAAC 87327
 QY 1799 TGTATGCGATACCACTTACTCTGCAAGAAATGTCATTAATTAATAATCAAAATTA 1858
 Db 87326 TGTATGCGAATCTTACTTACTCTGCAAGAAATGTCATTAATTAATAATCAAAATTA 87267
 QY 1859 TAGATGTTGGTGGTCTGTGTGATTAAGAAACATTTTACCTGCTGTGGAAATGA 1918
 Db 87266 TAGATGTTGGTGGTGTGATGAGTAAGCAAGAAACATTTTA-TGTGCTGTGGAAATGA 87208
 QY 1919 ACTTGGCAACCACTATGGAACCAAGTGGAAATTTCTTAAGAACTTAAGTGAATG 1978
 Db 87207 ACAAGTATGGCCACATGGAACCAAGTGGAAATTTCTTAAGAACTTAAGTGAATG 87148
 QY 1979 ACCATTTGATCCAGCAATCCCATTA----- 2003
 Db 87147 ACCATTTGATCCAGCAATCCCATTA----- 87088
 QY 2004 ----- 2003
 Db 87087 CGAAGAAATCTTGACACACGTTTATAGTACCAATTTGCAATTTGCAATTAAGTGA 87028
 QY 2004 -----AATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2034
 Db 87027 ACCAACCAATATCCATCAATCAACAGTGAATTAATTAATTAATTAATTAATTAATTA 86968
 QY 2035 TGAATATCACTCAGCATTAATAAAGATTAATGATGATTCACAGCAATCTAGATG 2094
 Db 86967 TGAATATCACTCAGCATTAATAAAGATTAATGATGATTCACAGCAATCTAGATG 86908
 QY 2095 AATTGAGACCTTATTTAAGTGGGTAATCTGAGATGAAACCAACATCATATGT 2154
 Db 86907 GATTGAGACCAATTAATTAAGTGAATGATCTGAGAAATGAAACCAACCTTATATGT 86848
 QY 2155 TCTCATTAATGATGGGCTAAGCTGTGAGGACAGAAAGCAT-AGAATGATTAATGA 2213
 Db 86847 TCTCATTAATGATGGGCTAAGCTGTGAGGATGAAAGCATTAAGTGAATGACATATG 86788
 QY 2214 ACTCTGGGACT-TGAGGGGAAGATGGAAGAGCGAGGATTAAGACTACAAATG 2272

Db 86787 ACTTGGAGACTCGGGGGAAAGGCTGGAGAAAGAGTGAGTTAAAGACTTACAAATTG 86728
Qy 2273 GGTACACTGTACACTGCTCAGGTGATGGGTGCACCAAAATCTCAGAAAT-TACCACTAAA 2331
Db 86727 GGTTCAGTGTATATATGCGCGGTGATGGGTGCACCAAAATCTCACAATGAACCACTTAA 86668
Qy 2332 GAACCTTATCCATGAGAGCAACACCACTGTTCCCAAAATCCCAATGAATTAATAATA 2391
Db 86667 GAACCTTACTATGTACCAATATACCACTGTTCCGATACCTATGGAATAAAAAACAA 86608
Qy 2392 TAATATATAA 2401
Db 86607 AAAACAAACA 86598
RESULT 14
AAK68992
ID AAK68992 standard; DNA; 42738 BP.
AC AAK68992;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23804.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KM cytosolic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN W0200157182-A2.
PD 09-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0214866P.
PR 28-JUN-2000; 2000US-0219467P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218230P.
PR 26-JUL-2000; 2000US-0230963P.
PR 26-JUL-2000; 2000US-0230964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241021P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241788P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.

[illegible]

QY	2154	TTCTCACTTACCAAGTGGGGGGCTTAAGCTGTGAGGACACCAAGGCA-TAGAAATGATTAATG	221
Db	20966	TTCTCACTTACATTAAGTGGAGGAGCTTAAGCTCTGAGGATTTAAAGCGAAGAAATGATTAACAAG	210
QY	2213	AACTCTGGGGAGCTTGAAGGGAGATGGAAGAGAGGC-GAGGATTAAGAAAGCTACACAAT	222
Db	21086	GGGTTCAGGTGAAGTGAACCTGCTGGGATGATGATGTCACCAAAATCTCACAAATACCAATTA	211
QY	2272	GGGTTCAGGTGAAGTGAACCTGCTGGGATGATGATGTCACCAAAATCTCACAAATACCAATTA	223
Db	21066	GGGTTCAGGTGAAGTGAACCTGCTGGGATGATGATGTCACCAAAATCTCACAAATACCAATTA	211
QY	2332	GAACCTTATCATGGAAGCAACCAACCTGTTCCCAAAATCCCAATGANA	2382
Db	21146	GAACTTATCATGATTAACCAAAATCAACCTGTTCCCAACCAAAATCAATGANA	21196
RESULT 15			
ID	ABK83575	standard; cDNA; 136284 BP.	
AC	ABK83575;		
XX			
DT	14-AUG-2002	(first entry)	
XX			
DE	Human cDNA differentially expressed in granulocytic cells #146.		
XX			
KM	Human; ss; granulocytic cell; DNA chip; bacterial infection;		
KM	viral infection; parasitic infection; protozoal infection;		
KM	fungal infection; sterile inflammation; disease; psoriasis;		
KM	rheumatoid arthritis; glomerulonephritis; asthma; chronic;		
KM	cardiac reperfusion injury; renal reperfusion injury; ARDS;		
KM	adult respiratory distress syndrome; inflammatory bowel disease;		
KM	Crohn's disease; ulcerative colitis; periodontal disease;		
KM	granulocyte activation; chronic inflammation; allergy.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200228999-A2.		
XX			
PD	11-APR-2002.		
XX			
PF	03-OCT-2001; 2001WO-US030821.		
XX			
PR	03-OCT-2000; 2000US-0237189P.		
XX			
PA	(GENE-) GENE LOGIC INC.		
XX			
PI	Beazer-Barclay Y, Weisman SM, Yamaga S, Vockley J;		
DR	WPI; 2002-435328/46.		
XX			
PT	Detecting granulocyte activation by detecting differential expression of		
XX	genes associated with granulocyte activation, which serves as diagnostic		
PT	markers that is useful for monitoring disease states and drug toxicity.		
XX			
PS	Claim 1; SEQ ID NO 146; 114pp; English.		
XX			
CC	The invention relates to detecting (M1) granulocyte (GC) activation		
CC	(GCA), by detecting the level of expression of gene(s) (Gs) identified by		
CC	DNA chip analysis as given in the specification, and comparing the		
CC	expression level to an expression level in an unactivated GC, where		
CC	differential expression of Gs is indicative of GCA. Also included are		
CC	modulating (M2) GA by contacting GC with an agent that alters the		
CC	expression of at least one gene in Gs; (2) screening (M3) for an agent		
CC	capable of modulating GCA or an inflammation (especially chronic) in a		
CC	tissue, an allergic response in a subject, exposure of a subject to a		
CC	pathogen or sterile inflammatory disease using the gene expression		
CC	profile; (3) detecting (M4) an inflammation (especially chronic) in a		
CC	tissue, an allergic response (M4) an inflammation (especially chronic) in a		
CC	pathogen or sterile inflammatory disease, by detecting the level of		
CC	expression in a sample of the tissue of gene(s) from Gs, where the level		
CC	of expression of the gene is indicative of inflammation; (4) treating		

(M5) an inflammation (especially chronic) or in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease, by contacting a tissue having inflammation with an agent that modulates the expression of gene(s) from Gs in the tissue. M1 is useful for detecting GCA; M2 is useful for modulating GAI; M3 is useful for screening an agent capable of modulating GCA preferably in an inflammation in a tissue; M4 is useful for detecting an inflammation (especially chronic) in a tissue, an allergic response in a subject, exposure of a subject to a pathogen or sterile inflammatory disease (e.g. psoriasis, rheumatoid arthritis, glomerulonephritis, asthma, thrombosis, cardiac reperfusion injury, renal reperfusion injury, ARDS, adult respiratory distress syndrome, inflammatory bowel disease, Crohn's disease, ulcerative colitis, periodontal disease, also bacterial infection, viral infection, parasitic infection, protozoal infection, fungal infection and M5 is useful for treating one of the above conditions. The present sequence represents a gene differentially expressed in granulocytes. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in the electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 136284 BP; 41693 A; 28668 C; 28000 G; 37923 T; 0 U; 0 Other;

Query Match 15.6%; Score 1346; DB 6; Length 136284;
Best Local Similarity 76.6%; Pred. No. 4,1e-302;
Matches 1985; Conservative 0; Mismatches 390; Indels 215; Gaps 19;

QY 1 TGTATGAAGCCAAATGCTACTTTATATCCAAACGAGAAAGATAT--CAAAAAGAAA 58
DB 83890 TCTATGAAGCCAGTGTCACTTATATCCAAACGAGAAAGGCAATACCAAAAAACAA 83949
QY 59 ACTATGACCAATACCACTGATGATATATCATGCAAGAAATCCCAACAAATCTAGCTA 118
DB 83950 ACGATGACCAATACCTGATGATATATCATGATGCTGATCTTAAACAAATCTAGCTA 84009
QY 119 ACCCAATCCAAAGCATATCAAGAAATATCCACATGCTCAAGTGGTTTCATACAG 178
DB 84010 ACTGATCCAAAGCATATCAAGAAATATCCACATGATCAAGTGGTTTCATACAG 84069
QY 179 GGGTCAGG-ATAGGTTAACTACCAAGTCAATTAATGATACATCATATAACGAA 237
DB 84070 GGATCAGAGAAATGTTAAATACCAAGTAAATCAAGTAAACCATATAGCGAA 84129
QY 238 TTAATAAC-AAATAACATGATCATCTCAATGATGCTGAAAGCATTTGCAAAATC 296
DB 84130 TTAATAACAAATAATACATGATCATCTCAATGATGCAAGAAAGATTTCAACAAATC 84189
QY 297 TAACTTTCTTTATGATTTAAACCTTCAGCAAAATGCAATAGAAAGACATACCTT 353
DB 84190 CAGCATCCCTTTATGATTTAAATCTCAGCAAAATGCAATAGCAACACACCTTGT 84249
QY 354 -----AATGATTAATAAGCATATATGACGAGCCACAGCAAAACATTAATCTGANT 405
DB 84250 TGTATTAATAATGATTAATAAGCATATGACGAGCCACAGCAAAACATTAATCTGANT 84309
QY 406 GGGAAAGTGAATAATGTC-----CTGAGACTGGAAC 442
DB 84310 GGGAAAGTGAATAATGTCCTCTGAGAACTGGAAGCATTCCTCTGAGAACTGGAAC 84369
QY 443 AAGACAGAGATG-CTACTTTCACTCACTTATTCATATGATGAGTGAAGTTTTCAGAG 501
DB 84370 AAGACAGAGATGCTCACTCTCACTCTTCAACAGAGTCTGGAATCTCGAGCAG 84429
QY 502 AGCAATCAGACAGAGAAAGAAATCAAGGCACTCCCAATCAATTAAGAGAGTCAAACT 561
DB 84430 AGCAATCAGACAGAGAAAGAAATCAAGGCACTCCCAATCAATTAAGAGAGTCAAACT 84489
QY 562 GTCCCTGTCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 621
DB 84490 GTCACTGTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 84549
QY 622 GCTCTAGAACTGATATCATTAATTCAGTAAAGTTTCAGATACAACTTAATATGACAA 681
DB 622 GCTCTAGAACTGATATCATTAATTCAGTAAAGTTTCAGATACAACTTAATATGACAA 681

DB 84550 GCTCTAGAACTGATATCATTAATTCAGTAAAGTTTCAGATACAACTTAATATGACAA 84609
QY 682 ATCAGTAGACCTGCTATATACCAACAGTACCAAGCTGAGATCAATCAAGATCTCAA 741
DB 84610 ATCAGTAGACCTGCTATATACCAACAGTACCAAGCTGAGATCAATCAAGATCTCAA 84667
QY 742 CACTTTTCAATATAGCTGTAAA-----AAATACTTAAAGATATTTCTTACCAAGAG 793
DB 84668 CCTTTTCAATATAGCTGTAAA-----AAATACTTAAAGATATTTCTTACCAAGAG 84727
QY 794 GTGAGAGACTTTCAGAGAAATCTCAAAACAGACT---GACTCATGATGACACA 849
DB 84728 TCAAAAGACTTTCAGAGAAATCTCAAAACAGACT---GACTCATGATGACACA 84787
QY 850 AACAAATGGAACACATCCCATGCTCATGAGTGGTGAAGTGAATGATGATGATGATGAT 909
DB 84788 AACAAATGGAACACATCCCATGCTCATGAGTGGTGAAGTGAATGATGATGATGATGAT 84847
QY 910 ATATGCAAAAGCAATCTCAAGATGCAATGCAATTCACCAACAAATATCATCATATTC 969
DB 84848 ACATGCAAAAGCAATCTCAAGATGCAATGCAATTCACCAACAAATATCATCATATTC 84907
QY 970 TTCAGAACTAGAAAAAAACATTTTAAATTCATATGAAACAAACAAAAA 1029
DB 84908 TTCAGAACTAG--AAAAACATTTCTGAATTTATATGAAAC-----CAA 84951
QY 1030 AAAAAAACCCTGATGAGCAAGCAAGTCTTACCAAAAGAAAGAAATCTGAGGCTACCA 1089
DB 84952 AAGAGGCAAGCAATGAGCAAGCAAGTCTTACCAAAAGAAAGAAATCTGAGGCTACCA 85011
QY 1090 TTACCATCTTCAAACTATATCTCAAGGCTATATACCAAAACATTCATGCACTGACAT 1149
DB 85012 CTACCTGATTTCAAACTCTATATAGGCTATATAGGCTATATAGGCTATATAGGCTAT 85071
QY 1150 AAAATGACATATAGCAAT-----GAAAAAGAGAGAAATCCAAATTAAGC 1201
DB 85072 AAAATGACATATAGCAAT-----GAAAAAGAGAGAAATCCAAATTAAGC 85131
QY 1202 CAATTAATATAGGCAAGTATTTTGAACAAAGCAAAACAAATTAAGTGGGAAAG 1261
DB 85132 CAATTAATATAGGCAAGTATTTTGAACAAAGCAAAACAAATTAAGTGGGAAAG 85190
QY 1262 ACATTTAGTTAACAATGATGCTGAGATTAATGCAACCAATGAGAAATGAAAC 1321
DB 85191 ACATTTAGTTAACAATGATGCTGAGATTAATGCAACCAATGAGAAATGAAAC 85250
QY 1322 TGGATCCCTGCTCTCACTTAATCAAAATTTGATCAAGTGCATCAAACTTAAT 1381
DB 85251 TGGAT--CTCATCTCTCACTTAATCAAAATTTGATCAAGTGCATCAAACTTAAT 85309
QY 1382 CTGAGACCTAAACCAATTAATTTCTGAGAGATTAATGCAAGAAATGCTTACGACTT 1441
DB 85310 CTGAGACCTAAACCAATTAATTTCTGAGAGATTAATGCAAGAAATGCTTACGACTT 85369
QY 1442 CACTTGGCAAGACTTCAATGAGCAAGAACTCAAAAGTAAATGCAACAAACAAATA 1501
DB 85370 GCGTTAGCAAGATTTCAATGAGCAAGAACTCAAAAGTAAATGCAACAAACAAATA 85429
QY 1502 AATAGATAGGACTTAATTAATTAATAAGCTTTTGGCAGCAAAACAAATCTTGGCAG 1561
DB 85430 AATAGATAGGACTTAATTAATTAATAAGCTTTTGGCAGCAAAACAAATCTTGGCAG 85488
QY 1562 GCAACAGCAACCCAGAGTGAAGAAATCTTCAAAACCTTAAGCATCTGATTAAGGA 1621
DB 85489 GCAACAGCAACCCAGAGTGAAGAAATCTTCAAAACCTTAAGCATCTGATTAAGGA 85548
QY 1622 CTAAATTCGAGATTCACAGAGAACTCAAAACAAATGAGCAAGAAAGCAAACTATCC 1681
DB 85549 CTAAATTCGAGATTCACAGAGAACTCAAAACAAATGAGCAAGAAAGCAAACTATCC 85606
QY 1682 ATGAAAGAGTGGCTAAGACATGAATTAAGCAATTTCAAAAGAGATATCAATGACC 1741
DB 85607 ATGAAAGAGTGGCTAAGATGAATTAAGCAATTTCAAAAGAGATATCAATGACC 85666

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 2, 2004, 00:25:44 ; Search time 692 Seconds

(without alignments)
8856.101 Million cell updates/sec

Title: US-09-867-570-3

Perfect score: 8622

Sequence: 1 tgcattgaagccatgctact.....gtccccaagccctacc 8622

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCUTUS.COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	1286.8	14.9 148567	US-09-801-876B-3	Sequence 3, Appli
2	1286.8	14.9 148567	US-10-254-869-3	Sequence 3, Appli
3	1106.8	12.8 32042	US-09-245-281-44	Sequence 44, Appli
4	1106.8	12.8 32042	US-09-340-620A-63	Sequence 63, Appli
5	1106.2	12.8 116592	US-09-818-512-3	Sequence 3, Appli
6	1069.4	12.4 193303	US-09-497-855A-37	Sequence 37, Appli
7	1069.4	12.4 193303	US-09-497-855A-44	Sequence 44, Appli
8	1069	12.4 40328	US-08-742-185-102	Sequence 102, App
9	1068.6	12.4 392000	US-10-027-983-11	Sequence 11, Appli
10	1064.6	12.3 168174	US-10-071-411A-63	Sequence 63, Appli
11	1064.6	12.3 168174	US-10-071-411A-2	Sequence 2, Appli
12	1064.2	12.3 99916	US-09-816-095-3	Sequence 3, Appli
13	1058.6	12.3 168174	US-10-071-411A-63	Sequence 63, Appli
14	1058.6	12.3 168273	US-10-071-411A-2	Sequence 2, Appli
15	1054.4	12.2 786431	US-09-751-389-3	Sequence 3, Appli
16	1045.6	12.1 18596	US-09-318-448-11	Sequence 11, Appli
17	1045.6	12.1 18597	US-09-962-665-8	Sequence 8, Appli
18	1045.6	12.1 18597	US-09-962-665-8	Sequence 8, Appli
19	1045.6	12.1 18597	US-09-962-665-8	Sequence 8, Appli
20	1045.6	12.1 18597	US-09-962-665-8	Sequence 8, Appli
21	976.8	11.3 10607	US-08-078-090-3	Sequence 3, Appli
22	968.8	11.0 99916	US-09-816-095-3	Sequence 3, Appli
23	949.8	10.9 969	US-09-254-227A-4	Sequence 4, Appli
24	938.6	10.9 969	US-09-254-227A-6	Sequence 6, Appli
25	938	10.9 786431	US-09-751-389-3	Sequence 3, Appli
26	930.6	10.8 168575	US-09-426-290-1	Sequence 1, Appli
27	873.6	10.1 392000	US-10-027-983-11	Sequence 11, Appli

ALIGNMENTS

C	28	863.6	10.0 197496	4	US-09-877-177A-10	Sequence 10, Appli
	29	836.2	9.7 969	4	US-09-254-227A-12	Sequence 12, Appli
	30	833	9.7 969	4	US-09-254-227A-8	Sequence 8, Appli
	31	829.8	9.6 969	4	US-09-254-227A-14	Sequence 14, Appli
	32	826.6	9.6 969	4	US-09-254-227A-10	Sequence 10, Appli
	33	821	9.5 29485	4	US-09-785-381-6	Sequence 6, Appli
	34	802.8	9.3 38844	3	US-09-734-675-3	Sequence 3, Appli
	35	800.8	9.3 29604	3	US-08-781-891-207	Sequence 207, App
	36	800.8	9.3 29604	3	US-09-618-166-207	Sequence 207, App
	37	761.6	8.8 90541	4	US-09-759-359A-3	Sequence 3, Appli
	38	761.6	8.8 90541	4	US-10-207-973-3	Sequence 3, Appli
	39	709.4	8.2 38059	4	US-09-328-925-4	Sequence 4, Appli
	40	706.2	8.2 80595	3	US-09-078-294-4	Sequence 3, Appli
	41	704.6	8.2 80246	3	US-09-078-294-4	Sequence 3, Appli
	42	672.8	7.8 118067	4	US-09-457-855A-32	Sequence 32, Appli
	43	657.2	7.6 15418	4	US-09-783-203-1	Sequence 1, Appli
	44	657.2	7.6 15418	4	US-09-994-427A-1	Sequence 1, Appli
	45	657.2	7.6 15418	4	US-09-244-438-1	Sequence 1, Appli

RESULT 1	US-09-801-876B-3	Sequence 3, Application US/09801876B
Patent No. 6492155	GENERAL INFORMATION:	
APPLICANT: YE, Jane et al	TITLE OR INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC	
TITLE OR INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES		
FILE REFERENCE: CLO01160		
CURRENT FILING DATE: 2001-03-09		
NUMBER OF SEQ. ID NOS: 8		
SOFTWARE: FASTSEQ for Windows Version 4.0		
SEQ ID NO 3		
LENGTH: 148567		
TYPE: DNA		
ORGANISM: Human		
FEATURE:		
NAME/KEY: misc_feature		
LOCATION: (1)..(148567)		
OTHER INFORMATION: n = A,T,C or G		
US-09-801-876B-3		
Query Match	14.9%; Score 1286.8; DB 4; Length 148567;	
Best Local Similarity	76.3%; Pred. No. 0;	
Matches 1972; Conservative	0; Mismatches 392; Indels 220; Gaps 23;	
QY	3	TATGAGCCATGTCATCTTAATACCAAAACGAGAAAGATATA--CAAAAAGAAAC 60
DB	45746	TATGAGCCATGTCATCTTAATACCAAAACGAGAAAGATATA--CAAAAAGAAAC 45805
QY	61	TATGAGCCATGTCATCTTAATACCAAAACGAGAAAGATATA--CAAAAAGAAAC 120
DB	45806	TGCGAGCAATATCATGATGAGACAGATGCTGAATCTTAACAAAATAGTGTAC 45865
QY	121	CCATTCACAGCATATCAAGAAATATCCACATGTCAGTGGGTTTATACCGAGG 180
DB	45866	TGAATCCACAGCATATCAAGAAATATCCACATGTCAGTGGGTTTATACCGAGG 45925
QY	181	GTCGAGG-ATAGGTTAATACATCAAGTCAATTAATGATATGATACATTAACAGAAATT 239
DB	45926	ATGCGAATATGCTTAATACATCAAGTCAATTAATGATATGATACATTAACAGAAATT 45985
QY	240	AAAAACCAATCATGATCATCTCAATAGATGCTGAAGAAAGATTGACAAATGTAA 299
DB	45986	TTTAA-AAAATCATGATCATCTCAATAGATGCTGAAGAAAGATTGACAAATGTAA 46044
QY	300	CATTCTTTGATTAACCTTCAGAAATGCAATAGAAAGATATA--CCTTAATG 357

Db 46045 CATCCCTTTATGATTAACCCCTCAGCAAAATCAGCATACAAAGGACATAGCCCTTAATG 46104
 Qy 358 TAAATAAAGCCATATATAGGAGACCCAGCAAAACATTAATCTGAATGGGAAAAAGTTGA 417
 Db 46105 TAAATAAAGCCATATATAGCAAAACCCAGCCAAACATTAACCTGAAC----- 46151
 Qy 418 AAACATGTCCTCCAGAACTGGAAACAAGCAAGATG-CTACTTCCACCACTTCTAATCA 476
 Db 46152 --ACATCTCCTTGAGAACCAAGATGAGACAGATGAGCCACTCTCACTGCTCTTCA 46209
 Qy 477 ACATAGTAGTGAAGTTTATGCCAGAGCAATCAGACAGAAAGAAATCAAGGACCC 536
 Db 46210 ATGTAGTACTGGAAGTCTAGCCAGAGCAATAGACAAAGAAATTAAGTCACT 46269
 Qy 537 AAATCAATTAAGAGAAAGTCAAACTGCTCTGTTCACTGATGATGATGTTATACCTAG 596
 Db 46270 AAATCAATTAAGAGAAAGTCAAACTGCTCTGTTCACTGATGATGATGTTATACCTAG 596
 Qy 597 AAACCCCTAAAGCTCATCCAGAAAGCTCCTAGAACATGATACATTAATTCAGTAAAGTT 656
 Db 46330 AAACCCCTAAAGCTCCTCAGAAAGCTCCTAGAACATGATACATTAATTCAGTAAAGTT 656
 Qy 657 CAGGATCAAACTAAATGTAACAAATCAGTACACTGCTATACACCAACAGTACCAAG 716
 Db 46390 CCGGATCAAACTAAATGTAACAAATCAGTACACTGCTATACACCAACAGTACCAAG 716
 Qy 717 CTGAAATCAATCAAGAACTCAAACTTTTACATAGCTGMAA----- 763
 Db 46450 TAGAGAACCAATCAAGAACTCAATCCCTTTTACATAGCTGMAA----- 763
 Qy 764 -----AAATCTTAAGATATTTCTTACCAGAG 793
 Db 46510 ACAAGCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 793
 Qy 794 GTGAAGGACCTCTACAA-GGAAACCTACAAACAGCT---GACATCATAGTACAC 848
 Db 46570 TAGAAAGACCTCTACAAAGGAAATTAACAAACAGCTGGAAGATCATATAGTACAC 848
 Qy 849 AAACAAGTGAACACATCCCATGCTCATAGAGGGGAGAAATCAATTTGTTG-AAAAAGA 907
 Db 46630 AAACAAGTGAACACATCCCATGCTCATAGAGGGGAGAAATCAATTTGTTGAAAAATA 907
 Qy 908 CCATATGCGCAAAAGCAATCTACAGATTCATGCAATTCACCAAAATTCATCATCAT 967
 Db 46690 CCATATGCGCAAAAGCAATCTACAGATTCATGCAATTCATGCAATTCATCATCAT 967
 Qy 968 TCTTCAAGAACTAGAAAAAAACAATTTCAAAATTCATATGAAACAACAACAAAAAA 1027
 Db 46750 TCTTCAAGAACTAG-AAAAACAATTTCAAAATTCATATGAAAC-----C 46794
 Qy 1028 AAAAAAAGCCGCAATAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1087
 Db 46795 AAAAAAAGCCGCAATAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1087
 Qy 1088 CATTAACCACTCTTCAAACTATACAGAGGCTAATTCACCAAAACATCATGCACTGAC 1147
 Db 46855 CATTAACCACTCTTCAAACTATACAGAGGCTAATTCACCAAAACATCATGCACTGAC 1147
 Qy 1148 ATAAAACTAGGACATAGCAATGAAAGAAAGAAATTCAGAAATTAAGCCAAATA 1207
 Db 46915 AAAAAAATAGGACATAGCAATGAAAGAAATGAAAGAAATTAAGCCAAATA 1207
 Qy 1208 ATTAATGCAATGATTTTGAACAAGCAAAACAATTAAGTGGGAAAGCAATTC 1267
 Db 46975 CTTAACGCACTGATCTTGAATGAAGCAAAATGAATTAAGTGGGAAAGCAATTC 1267
 Qy 1268 TAGATTAACAATGCTGATGATTAATGCGCAAGCAATGGAAGATGAATGCACTGATC 1327
 Db 47035 TTTTCAACAATGCTGATGATTAATGCGCAAGCAATGGAAGATGAATGCACTGATC 1327
 Qy 1328 CCTTGTCTCTCACTTAATCAAAATTTGATACAGATGATCAAGAATTAATCTGAGA 1387
 Db 47094 CCTTGTCTCTCACTTAATCAAAATTTGATACAGATGATCAAGAATTAATCTGAGA 1387
 Qy 47094 CCTTGTCTCTCACTTAATCAAAATTTGATACAGATGATCAAGAATTAATCTGAGA 1387

Qy 1388 CCTAAACCATAAATTTCTAGAGTAAATCATCAGAAAAATGCTCTGACATTCATTA 1447
 Db 47154 CTTGAATCTTAATTAATTTTAAAGATTAATTTGAAAAACCTTTCTGACATTTGGCTTA 47213
 Qy 1448 GCGAAAGCTTATGCGCAAGAACCCAAAGTAAATGCAACAAAAAATAATTAATGA 1507
 Db 47214 AGCAAGGTTTATGACCAAGAACCCAAAGTAAATGCAATTAATTAATTAATG 47273
 Qy 1508 TAGGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1567
 Db 47274 TGGTACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1567
 Qy 1568 AGCAACCCAGAGTGAAGAAATTTCTCAAACTAAGCACTGACCTAAGCACTAATTA 1627
 Db 47323 AAACGCCACAGATGGAAGAAATTTCTCAAACTAAGCACTGACCTAAGCACTAATTA 1627
 Qy 1628 TCCGAATCC-ACAAGAACTCAAAACAATCAGCAAGAAAGCAAAACAATCCCATGA 1686
 Db 47383 TCCGAATCTCAATGAATCAAGTCAAGTAAATCAGTAAAGAA-----AAACAATCTATCA 47438
 Qy 1687 AGAGTGGCTTAAGGACATGAATGACAAATTCACAAAGAAATTAATTAATTAATTAAT 1746
 Db 47439 AAAGTGGCTTAAGGACATGAATGACAAATTCACAAAGAAATTAATTAATTAATTAAT 1746
 Qy 1747 AC--AGAAAAATGCTTAATCACTAATGATTAAGGAAATGTAATCAACTGTAAT 1804
 Db 47499 ACATGAAGAAATGCTCAACATCACTAATGATTAAGGAAATGTAATCAACTGTAAT 1804
 Qy 1805 GCGATACCACTTACTCTGCAAGAAATGCTAAT--TTAAATCTTAATTAATTAATGA 1861
 Db 47559 GTGATTCACCTTACTCTGCAAGAAATGCTAATTAATTAATTAATTAATTAATTAATGA 1861
 Qy 1862 ATGTTGGTGGTCTGCTGATTAAGAAACAATTT-TACATGCTGGTGGAAATGAATGA 1920
 Db 47619 ATGTTGGTGGTCTGCTGATTAAGAAACAATTT-TACATGCTGGTGGAAATGAATGA 1920
 Qy 1921 TTGCGCAACCACTATGCAAGAAACAGTGGAAATTTCTTAAGAACTAAAGTGAATGAC 1980
 Db 47679 TAGTACAGCACTATGGAACAGTGGAAATTTCTTAAGAACTAAAGTGAATGAC 1980
 Qy 1981 CATTTGATCAGCAATGCC----- 1999
 Db 47739 CATTTGATCAGCAATGCCCTCTACTGGTATCTACAGAGAAATTAATGATTAATCA 1999
 Qy 2000 ----- 1999
 Db 47799 AAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1999
 Qy 2000 -----ATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2052
 Db 47859 ATCAATGATGATTAAGAAACCTGCTATATGATTAATTAATTAATTAATTAATTAAT 2052
 Qy 2053 TAAAAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2112
 Db 47919 TAAAAAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2112
 Qy 2113 TAAGTGGGTAATCAAGAAATGAAACCAACATCAATGTTTCACTTAACAATGGGG 2172
 Db 47979 TAAGTGGGTAATCAAGAAATGAAACCAACATCAATGTTTCACTTAACAATGGGG 2172
 Qy 2173 GCTAAGCTGAGAGCAAGAGGAT-AGATGATTAATGAATCTGAGGAGCTGAAGG 2231
 Db 48039 GCTAAGCTGAGAGCAAGAGGAT-AGATGATTAATGAATCTGAGGAGCTGAAGG 2231
 Qy 2232 GAAGATGAAGAGAGGCGAGGATTAAGATTAATGAATCAATGAGTACAGTCTGCT 2291
 Db 48099 GAAGATGAAGAGAGGCGAGGATTAAGATTAATGAATCAATGAGTACAGTCTGCT 2291
 Qy 2292 AGGTATGAGTGAACCAAAATCTCAGAAATTAATTAAGAACTTAATCAATGAGGAA 2351
 Db 48159 AGGTATGAGTGAACCAAAATCTCAGAAATTAATTAAGAACTTAATCAATGAGGAA 2351
 Qy 48217 AGGTATGAGTGAACCAAAATCTCAGAAATTAATTAAGAACTTAATCAATGAGGAA 48217

Db 13932 CAGAGACCAATATAGGTGAATCCCATTTGCTTCAAGAAATTAATATCC 13873
 Qy 771 TAAAGATATTTCTTACCAAGAGGTGAAGACCTCTACAGAGAAATCAACAAACAGAC 830
 Db 13872 TAGGAATTCGAATTAACAAGGAGTGAAGACCTCTTCAAGAGAAATCAACAAACAGAC 13813
 Qy 831 T----GACATCATGATGACACAAAGTGGAAACACATCCCTGCTCATGATGGGTA 886
 Db 13812 TCAAGAAATTAAGAGACCAACAAATGGAAGAACTTCCATGCTCATGGGTAGGAA 13753
 Qy 887 GAATCATATTTGTGAAATGACATATTTGCCAAAGAAATCTACAAAGTTCAATGCAATTC 946
 Db 13752 GAATCATATTTGTGAAATGACATATTTGCCAAAGAAATCTACAAAGTTCAATGCAATTC 13693
 Qy 947 CCACCAAAATATCATCATCTTCTTCAAGAAATGAAAGAAAGAAATCTTAAATTCATA 1006
 Db 13692 CCATCAAGCTTCAATGATCTTTCTTCAAGAAATGAAAGAAAGAAATCTTAAATTCATA 13634
 Qy 1007 TGGAAACAACAACCAAAAAAAAAAAAAAAAACCCGCATGCGCAAGCAAGCTTATGCAAAA 1066
 Db 13633 TGGAAAC-----CAAAAAAGAGCCGCGCATGCGCAAGCTTATGCAAGCAAAA 13588
 Qy 1067 AGAACAATCTGAGGACATCATTAACCATCTTCAAACTATCTACAAAGGCTATATCA 1126
 Db 13587 AGAACAATCTGAGGACATCATTAACCATCTTCAAACTATCTACAAAGGCTATATCA 13528
 Qy 1127 CCAAAACATCATGACATGACATTAACCTAGGACATGACCAATGAAAGAAAGAGAGA 1186
 Db 13527 CCAAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 13468
 Qy 1187 ATCCAGAAATTAAGCCAAATTAATTAAGCCAACTGATTTTGAACAAACAAACAAACA 1246
 Db 13467 CCTCAGAAATTAAGCCAAATTAATTAAGCCAACTGATTTTGAACAAACCTGAGAAAA 13408
 Qy 1247 TAAAGTGGGAAAGACATTTAGTTAACAAAATGATGATGATGATGATGATGATGATGAT 1306
 Db 13407 AGCAATGGGAAAGATTTCCCTATTTAATTAATTAATTAATTAATTAATTAATTAAT 13348
 Qy 1307 GTGAGAAATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 13366
 Db 13347 GTGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 13289
 Qy 1367 ATCAAAAGCTTAATCTGAGACCTTAAACCTTAAATTTTGAAGATTAATCAAGAAA 1426
 Db 13288 ATTAAGATTTAAAGCTTAAACCTTAAACCTTAAAGATTAATCAAGAAA 13229
 Qy 1427 ATGCTTGAACATCTTGAAGCAAGCTTCAATGCGCAAGAAACCAAAATGATTAATGCA 1486
 Db 13228 ACCATTCAGACATAGGCGTGGGCAAGGATTCATGTCACAAACCAAAAGCAATGGCA 13169
 Qy 1487 ACAAAACAAAAATTAATAGATAGATTAATTAATTAATTAATTAATTAATTAATTAAT 1546
 Db 13168 ACAAAACAAAAATTAATAGATAGATTAATTAATTAATTAATTAATTAATTAATTAAT 13109
 Qy 1547 ACAATCTTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1606
 Db 13108 GAAATCTACATCAAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 13049
 Qy 1607 GCATCTGAATGAGATTAATTCGGGAATCCCAAGAAATCAAAATCAAGAAAGAG 1666
 Db 13048 TCATCTGAACAAAGGCTTAATTAATCAAGATCTACATGATCAAAATTTTACAGAAA 12989
 Qy 1667 AAAGCAAAATCCATGAAAGAGTGGCTAAGCAATGATGATGATGATGATGATGATGATGAT 1726
 Db 12988 AAAGCAAAATCCATGAAAGAGTGGCTAAGCAATGATGATGATGATGATGATGATGATGAT 12929
 Qy 1727 GATTAACAATGCGCAACA--ACAGGAAATGCTTAACATCACTAATGATGAGGAA 1784
 Db 12928 GACATTTATGAGCAAAACCAATGAAAGATGCTCATCATGCGGCATCAAGAAA 12869
 Qy 1785 ATGTAATCAACACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1844

Db 12868 ATGCAATCAAAACCACTATGATGATATCTCAGACAGTATGATGATGATGATGATGAT 12809
 Qy 1845 AAATCTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1904
 Db 12808 AAGTCAGAAACAGCAGTCTGAGAGATGCGGAAATGAGAAACATTTTACACTGT 12749
 Qy 1905 TGGTGGGAATGTAATCTTGGGCAACCACTATGAAACAGTGGAAATTTCTTAAGGAA 1964
 Db 12748 TGGTGGGACGTAACTAGTTCACCAATGGAAGTCAAGTGGCAATTCCTCAGGGAT 12689
 Qy 1965 CTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2024
 Db 12688 CTAGATCTAGAAATACATTTGACCGACCATCCCATTT-ACGGGTATATACCAATGAA 12630
 Qy 2025 TTATATACATG 2036
 Db 12629 GTATTAATCATG 12618

RESULT 4
 US-09-340-620A-63/c
 ; Sequence 63, Application US/09340620A
 ; Patent No. 6482933
 ; GENERAL INFORMATION:
 ; APPLICANT: Bertin, John
 ; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
 ; FILE REFERENCE: 07334-12401
 ; CURRENT APPLICATION NUMBER: US/09/340,620A
 ; PRIOR FILING DATE: 1999-06-28
 ; PRIOR APPLICATION NUMBER: US 09/245,281
 ; PRIOR FILING DATE: 1999-02-05
 ; PRIOR APPLICATION NUMBER: US 09/207,359
 ; PRIOR FILING DATE: 1998-12-08
 ; PRIOR APPLICATION NUMBER: US 09/099,041
 ; PRIOR FILING DATE: 1998-06-17
 ; PRIOR APPLICATION NUMBER: US 09/019,942
 ; NUMBER OF SEQ ID NOS: 71
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 63
 ; LENGTH: 32042
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-340-620A-63

Query Match 12.8%; Score 1106.8; DB 4; Length 32042;
 Best Local Similarity 74.7%; Pred. No. 1.7e-271;
 Matches 1532; Conservative 0; Mismatches 487; Indels 33; Gaps 10;

Qy 1 TGTATGAAGCAATGCTTAAATACCAAAACCAAGAAAGATATA--CAAAAAAGAAA 58
 Db 14652 TTTATGAGGCGAGCATATCTGATACCAAGCCGCGCAGAGACACAAACCAAAAAAGGAA 14593
 Qy 59 ACTATGACAGTACCACTGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 118
 Db 14592 ATTTAGACCAATATCTTGTATGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 14533
 Qy 119 ACCCAATCAACAGCATATCAAGAAATTAATCCACATTTGCAAGTGGGTTTCAATACAG 178
 Db 14532 ACCGAATCAACAGCATATCAAGAAATTAATCCACATTTGCAAGTGGGTTTCAATACAG 14473
 Qy 179 GGGTGC-AGGATAGTGAATCAATCAAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 237
 Db 14472 GGAATGAGGCGGTGATTAATGCAAAATCAAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 14413
 Qy 238 TTAATAACAAATTAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 297
 Db 14412 CCAAGACAAACCAATGAT 14353
 Qy 298 AACATTTCTTATGATTAATTAATCTTCAAGCAAAATGACATAGAAAGCAATACCTTAATG 357
 Db 14352 AACAACTTATGATGATTAATTAATCTTCAAGCAAAATGACATAGAAAGCAATACCTTAATG 14293

Db 51038 TTTATGAGGCGACGATCATTTCTGATACCAAGCGAGGACGACACCAAAAAAGGA 50979
 Qy 59 ACTATAGACCAAGTACCACTGATGAATATACATGACGAATCCCAACAAATACTACTA 118
 Db 50978 ATTTTACCAAAATCTCTTGAATGAACTGCAATGCAATGCAATGCAATGCAATGCA 50919
 Qy 119 ACCCAATCCACAGCATATCAAGAAATATCCACATTTGCAAGTGGGTTCTATCCAG 178
 Db 50918 ACTGAATCCAGCAGCAGCATCAAAAAGCTTATCCACATGATCAGGTGGGCTTCATCCCTG 50859
 Qy 179 GGGTGC-AGGATAGGTTATACATACAAAGTAAATGATGATCATCATCAATAACAGAA 237
 Db 50858 GGATGCAAGGCTGGTTCAATATACCAAAATCAATTAATGATCCAGCATATTAACAGAG 50799
 Qy 238 TTAATAACAAAATCAGATGATCATCTCAATAGATGCTGAATAAAGCATTTGACAAATCT 297
 Db 50798 CCAAGACAAAACCAATGATTTATCTCAATAGATGCAAAAAGGCTTTGACAAATTC 50739
 Qy 298 AACATTTCTTATGATTAATAAAGCTTCAGCAAAATCGACATAGAAAAGACATACCTTAATG 357
 Db 50738 AACCAACCTTCATCTAATAAAGCTTCATTAATAGGATTTGATGGAGCATATCTCAAAA 50679
 Qy 358 TAAATAAAGGCATATATAGGAGGACCAAGCAAACTATATCTGAATGCGAAAAGTTGA 417
 Db 50678 TAAATAAGAGCATATATATAGCAAAACCAAGCAATATCATATGAAATGGGCAAAACCTGG 50619
 Qy 418 AACATTTCTTCCTGAGATGGAACCAAGACAGAGATG-CTACTTTACACTTCTATTTCA 476
 Db 50618 AAGATTTCTTTGAAAACCTGGACAGAGAGAGATGCTCTCTCTCTGCTCTATTTCA 50559
 Qy 477 ACATATGATGGAAGTTTATGACCAAGCATCAGACAGAAAGAAATCAAGGGCAACC 536
 Db 50558 ACATATGATGGAAGTTTATGACCAAGCATCAGACAGAAAGAAATCAAGGGATATTC 50499
 Qy 537 AAATCAATTAAGGAGAGTCAAACTGTCCCTGTCATGATGATGATGATGATGATGATGATG 596
 Db 50498 AAATGAGAAAGAGAGAGTCAAACTGTCCCTGTCATGATGATGATGATGATGATGATGATG 50439
 Qy 597 AAAACCTTAAGACTCATCCAGAAAGCTCTGAGATGATCATTAATTCAGTAAAGTTT 656
 Db 50438 AAAACCTTAAGACTCATCCAGAAAGCTCTGAGATGATCATTAATTCAGTAAAGTTT 50379
 Qy 657 CAGATTCAAATCTAATGATGACAAATGATGATGATGATGATGATGATGATGATGATGATG 716
 Db 50378 CAGATTCAAATCTAATGATGACAAATGATGATGATGATGATGATGATGATGATGATGATG 50319
 Qy 717 CTGGAATCAAAATGATGACAAATGATGATGATGATGATGATGATGATGATGATGATGATG 770
 Db 50318 CAGAGACCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 50259
 Qy 771 TAAAGATATTTCTTACCAAGAGGATGAGAGCTCTCAAGAGAACTCAACAAACAGCAGC 830
 Db 50258 TAGGAATCCAACTTACCAAGAGGATGAGAGCTCTCAAGAGAACTCAACAAACAGCAGC 50199
 Qy 831 T----GACATCATGATGACCAAAAGTGAAGCAATCCATGCTCATGATGATGATGATGATG 886
 Db 50198 TCAAGGAATTAAGAGAGCAAAACAAATGAGAAATTCATTCATGCTCATGAGGATGAGAA 50139
 Qy 887 GAATCATATTTGGAAGAAATGACATATTTGCCAAAGAGATCTCAAGTTCAATGCAATTC 946
 Db 50138 GAATCATATTTGGAAGAAATGACATATTTGCCAAAGAGATCTCAAGTTCAATGCAATTC 50079
 Qy 947 CCAACCAATATCATCATCTTCTTCAAGAGATGAGAAATCAATTTCAATTAATTCATA 1006
 Db 50078 CCAATCAAGCTACCAATGATCTTCTTCAAGAGATGAGAAATCAATTTCAATTAATTCATA 50020
 Qy 1007 TGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1066
 Db 50019 TGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 49974
 Qy 1067 AGAACAATCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1126

Db 49973 AGAACAAGCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 49914
 Qy 1127 CCAAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1186
 Db 49913 CCAAAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 49854
 Qy 1187 ATCCAGAAATTAAGCCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1246
 Db 49853 CCTCAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 49794
 Qy 1247 TAAAGTGGGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1306
 Db 49793 AGCAATGGGAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 49734
 Qy 1307 GTGGAAGATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1366
 Db 49733 GTGGAAGATGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 49675
 Qy 1367 ATCAAGATCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1426
 Db 49674 ATTAAGATTTAAACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 49615
 Qy 1427 ATGCTTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1486
 Db 49614 ACCATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 49555
 Qy 1487 ACAAAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1546
 Db 49554 ACAAAAACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 49495
 Qy 1547 ACAATCTTGAAGAGAAACCAAGCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1606
 Db 49494 GAAATCTCAATGAGAGAAACCAAGCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 49435
 Qy 1607 GCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1666
 Db 49434 TCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 49375
 Qy 1667 AAAGCAAAATCCCATGAAAGAGTGGCTTAAGGACATGATGATGATGATGATGATGATGATGATGATG 1726
 Db 49374 AAAGCAAAATCCCATGAAAGAGTGGCTTAAGGACATGATGATGATGATGATGATGATGATGATG 49315
 Qy 1727 GATATCAAAATGCGCAAAA-CAGAAAAATGCTTAATCAATCAATCAATCAATCAATCAATCAAT 1785
 Db 49314 GACATTTATGAGCAAAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 49255
 Qy 1786 TGTAAATCAACCTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1845
 Db 49254 TGTAAATCAACCTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 49195
 Qy 1846 AATCTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1905
 Db 49194 AGTACGAAACCAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 49135
 Qy 1906 GGTGGAAATGAACTTGGCGCAACCAATGAGAAACAGTGTGAATTTCTTAAGAAC 1965
 Db 49134 GGTGGAAATGAACTTGGCGCAACCAATGAGAAACAGTGTGAATTTCTTAAGAAC 49075
 Qy 1966 TAAAGATG 2025
 Db 49074 TAAAGATG 2006
 Qy 2026 TATATACCATG 2036
 Db 49015 TATATACCATG 49005

RESULT 6
 US-09-497-855A-37/c
 ; Sequence 37, Application US/09497855A
 ; Patent No. 6605432
 ; GENERAL INFORMATION:
 ; APPLICANT: Huang, Tim

```

; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
; FILE REFERENCE: UWO1523
; CURRENT APPLICATION NUMBER: US/09/497,855A
; CURRENT FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/120,592
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: 60/118,760
; PRIOR FILING DATE: 1999-02-05
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
; LENGTH: 193303
; TYPE: DNA
; ORGANISM: Homo sapiens;
US-09-497-855A-37

Query Match      12.4%; Score 1069.4; DB 4; Length 193303;
Best Local Similarity 74.1%; Pred. No. 1.9e-261;
Matches 1497; Conservative 0; Mismatches 486; Indels 36; Gaps 10;

QY      1 TGTATGAAGCCATGCTACTTTATATACCAAAACGAGAAAGATAT--CAAAAAGAA 58
DB      93923 TTTATGAGGCCAGCATCTCTATATCAAAAGCCTGGCAGAGACAAACCAAAAAGAGA 93864
QY      59 ACTATAGACCATGACCTGATGATATATACATGACAGAAATCCCAACAAATACTGCTA 118
DB      93863 ATTTTGACCAATATCTTGATGACATGATGACAAATCCCAATAAATCTGCA 93804
QY      119 ACCCAATCCAGACGATATCAAGAAATATCCACATTTGTCAAGTGGTTTCATACAG 178
DB      93803 ACCCAATCCAGACGATATCAAAAGCTTATTCACATGATCAAGTGGCTTATCCG 93744
QY      179 GGGTGC-AGGATAGGTTAACTATACACAGTCAATTAATGTGATPACATCAATAACGAA 237
DB      93743 GGATGCAAGGCTGTTCAATATACGCAATCAATTAATGTATCCAGCATATTAACGAA 93684
QY      238 TTTAAACCAAAATCATGATCATCTCATATAGATGTGAAAAGCATTTGACAAATCT 297
DB      93683 CCAAGACCAAAACCAATGATATCTCAATGATGAGAAAGCTTTTCACAAATTC 93624
QY      298 AACATTTCTTTATGATTTAAACCTTCAGCAAAATCCACATAGAAAGGACATACCTTATG 357
DB      93623 AACCAACCTTCAGATCAACACCTCTCAATTAATTAAGTATGATGGAACGATCTCAAA 93564
QY      358 TAATAAAGCCATATATGACGAGACCCACAGCAAAATTATATCTGATGGGAAAAGTTGA 417
DB      93563 TAATAAGACTATCTATGACAAACCCACACCAATATCATATGAAATGGCAAAAGCTGG 93504
QY      418 AAAACATTTGCCGAGAACTGGAACAGAAAGAGATG-CTACTTTCACCACTTCTATTC 476
DB      93503 AAGCATTTCCCTTGAAGAACTGGCAAGACAGGAGTCCCTCTCTCACTCTATTC 93444
QY      477 ACATATAGTGAAGTTTATGACGAGACATCAGACAAAGAAAGAAATCAAGGGCACCC 536
DB      93443 ACATATGTTGGAAGTTCTGGCCAGGGCAATTAGGACGAGAGAAATTAAGGTATTC 93384
QY      537 AAATCAATAAAGAGAGATCAAACTGTCTCTGTTCACTGATGATGATGATGATGATGATG 596
DB      93383 AAATTAAGAAAGAGAGATCAAAATTTGCTGTTGAGATGACATGATGATGATGATG 93324
QY      597 AAAACCCCTAAAGCTCATCCAGAAAGCTCTAGAACTGATACATTAATTCAGTAAATTT 656
DB      93323 AAAACCCATGCTCATCCAGCCAAATCTACTTAAGCTGATGAGCACTTCAGCAAGTCT 93264
QY      657 CAGGATACAAACTAAATGTACACAAATCAGTAGCACTGTATACCAACAGTAGACCAAG 716
DB      93263 CAGGATATTAATATCATATGTAACAAATCAAGCATTCATATACCA-----TAAACA 93208
QY      717 CTGAGATCAAAATCAAGAACTCAAACTTTTACATAGCTGTAA-----AAATACT 770
DB      93207 CAGAGAGCAAAATCATGATGAACTCCCATTCACAAATGCTTCAAAAGGAAATTAATATCC 93148
QY      771 TAAAGATATTTCTTACCAAGAGGATGAAGACTCTCTACAGAGAAATCTACAAACACAGC 830

```

```

DB      93147 TAGGAATCAACTTACAGAGGACATGAAGGGCTCTTCAAGAGAACTACAAACCACTGC 93088
QY      831 T-----GACATCTATGATGACACAAACAGTGGAAACATCCCTGCTCATGATGGTA 886
DB      93087 TCAAGAAATTAATAAGATGACAAAGAAATGAAGACTATCCATGCTCATAGGGTAGGAA 93028
QY      887 GAATCAATATTTGTAATAATGACCATATGCGCAAAAGCAATCTACAAAGTTCAATGCAATTC 946
DB      93027 GAATCAATATTTGTAATAATGCGCATATGCGCAATCTACAGATGTAATTTATGATTCAGTCCATCC 92968
QY      947 CCACCAAAATATCATCATCATTTCTTACAGAACTAGAAAACAAATTTAAATTCATA 1006
DB      92967 CCATCAAGGTACCAATGACTTTCTTACAGAAATGAAAACATCACTT--TAAAGTCATA 92909
QY      1007 TGGAAACACACCAAAAAAAAAAAAAAAAAAAAAAAACCCGATGCGCAAGCAAGACTTGGCAAA 1066
DB      92908 TGGAAAC-----AAAAAAAAAGCTGATCACCAAGTCAATCTTAAGCCATA 92863
QY      1067 AGAACAAATCTGAGGACATCACTTACCATCTTCAACTATATCTACAAAGCTATTAATCA 1126
DB      92862 AGAACAAAGCTGAGGACATCACTTACCATCTTCAACTATATCTACAAAGCTATACGTA 92803
QY      1127 CCAAAACATCATGCACTGACATTAATACTAGGACATAGACCAATGAAAAGAGAGA 1186
DB      92802 CCAAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 92743
QY      1187 ATCCGAATTAAGCCAAATATTTATGACCACTGATTTTGAACAAACCAAAACA 1246
DB      92742 CCTCAAAATTAATGCTGCTATCTCAAACTATCTATCTTGAACAACTGAGAAAACA 92683
QY      1247 TAAAGTGGGAAAAGCATTCTAGTTAACTAATGAGTGTGAGATTAATGGCAAGCAT 1306
DB      92682 AGCAATGGGAAAAGATTCCTATTTATTAATGATGATGATGATGATGATGATGATGATGATG 92623
QY      1307 GTGGAAGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1366
DB      92622 GTGGAAGATGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 92564
QY      1367 ATCAAGACTTAATTTGATGACCTTAACCAATTAATTTCTGAGATTAATCTGAGAA 1426
DB      92563 ATTAAGACTTAATTTGATGACCTTAACCAATTAATTTCTGAGATTAATCTGAGAA 92504
QY      1427 ATGCTCTAGACATTCATCTTAGCAAGACTCATGCGCAAGAACCCAAAGTAATGCA 1486
DB      92503 ACCATTCAGACATGAGCATGAGCAAGACTCATGCTTAACCAACCAAGCAATGCA 92444
QY      1487 ACNAAAACAAATAATTAATGATAGACTTAATTAATTAATTAATTAATTAATTAATTAAT 1546
DB      92443 ACNAAAACAAATAATTAATGATAGACTTAATTAATTAATTAATTAATTAATTAATTAAT 92384
QY      1547 ACATATATGACAGACAAACGACACCCAGGATGAGAAAATCTTCAAACTTA 1606
DB      92383 GAAACTTCACATCAGAGTGAACGACCACTTAATAATGAGAAAATTTTGGCAACTAC 92324
QY      1607 GCATCTGACTAAGGACTAATATCCGGAATCCCAAGAACTCAAACTCAAGCAAG 1666
DB      92323 TCATCTGACAAAGGCTGATATCCGGAATCTCAAGAACTCAAACTCAAGCAAG 92264
QY      1667 AAAGCAAACTCCATGGAAGAGTGGCTTAAGACATGATAGCAATTCCTCAAGAA 1726
DB      92263 AAAGCAAACTCCATGGAAGAGTGGCTTAAGACATGATAGCAATTCCTCAAGAA 92204
QY      1727 GATATACAAATGGCCAAACA--ACAGAAAATAATGTTAACTCACTAATGATTGGAA 1784
DB      92203 GACATTTATGACAGCCAAACAAACATGAAAATGCTCACTCACTGCGCCATCGAAG 92144
QY      1785 ATGTAATCAACCTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1844
DB      92143 ATGCAATCAAAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 92084
QY      1845 AAATCTAAATAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1904

```

Db 92083 AAGTCAGAAACCAAGTGTCTGAGAGATGTGAGAAATAGAAACCTTTACACTGT 92024
 QY 1905 TGTGGAAATGTAACTTGGCCAAACCACTATGAGAAAACAGTGTGAAATTTCTTAAGGA 1964
 Db 92023 TGTGGAACTGTAACTAGTTCAACCTTGTGGAAGTCAGTGTGGCCATTCCTCAGGAT 91864
 QY 1965 CTAAAGTAGTCCAGCTTTGTATCCAGCATCCCATTA 2003
 Db 91963 CTAGAACTAGAAATACCATTTGACCCAGCCATCCCATTA 91925

RESULT 7

US-09-497-855A-44/C
 ; Sequence 44, Application US/09497855A
 ; Patent No. 6605432
 ; GENERAL INFORMATION:
 ; APPLICANT: Huang, Tim
 ; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
 ; FILE REFERENCE: UMO1523
 ; CURRENT APPLICATION NUMBER: US/09/497, 855A
 ; CURRENT FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: 60/120,592
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: 60/118,760
 ; PRIOR FILING DATE: 1999-02-05
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 44
 ; LENGTH: 193303
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens;
 ; US-09-497-855A-44

Query Match 12.4%; Score 1069.4; DB 4; Length 193303;
 Best Local Similarity 74.1%; Pred. No. 1.9e-261;
 Matches 1497; Conservative 0; Mismatches 486; Indels 36; Gaps 10;

QY 1 TGTATGAAGCCATATGTCACTTTAATACCAAAACAGAAAGATATA--CAAAAAGAAA 58
 Db 93923 TTTATGAGGCGCAGCATCTCTTAATACCAAGCCTGGCAGAGACAAACAAAAGGA 93864
 QY 59 ACTATAGACCGATCCAGCTGATGAATATACATGAGAAATCCCAACAAATACTAGCTA 118
 Db 93863 ATTTGAGCCAAATCTCTGATGAACATGATGCAAAAATCTCAATAAATACTGGCA 93804
 QY 119 ACCCAATCCACAGCATATCAAGAAGATATCCACATTTCAAGTGGTTTCATACAG 178
 Db 93803 ACCAAATCCAGCAGCATCAAAAAGCTTATTCACATGATCAAGTGGCTTCATCCGTG 93744
 QY 179 GGGTGC-AGGATAGCTTAATACATACAGATCAATTAATGTGATACATCACTAAACAGAA 237
 Db 93743 GGAATGCAAGGCTGTTCATATATGCAAAATCAATTAATGTATATCAGCATTAACAGAA 93684
 QY 238 TTTAAAACAAAATCACATGATCATCTCAATAGATGTGAAAAAGCATTTGACAAAATCT 297
 Db 93683 CCAAGAGCAAAACCACTGATTAATCTCAATAGATGCAAAAAGCTTTTTCACAAAATTC 93624
 QY 298 AACATTTCTTATGATTTAAACCTTTCAGCAAAATGCAATAGAAAGCATACCTTAATG 357
 Db 93623 AAGAACCTTCAAGCTCAACCTCTCAATAAATTAGATTTGATGGACGTAATTCAAAA 93564
 QY 358 TAATTAAGGCATATATGACGACCCACAGCAAACTATTATAGATGGGAAAAAGTTGA 417
 Db 93563 TAATTAAGGCATATATGACCAACCCACAGCAATATCATCTGAATGGGCAAAAAGCTGG 93504
 QY 418 AACATTTGCTCCGAGATGGAACAGCAAGATG-CTACTTTCCACCACTTCTATTCA 476
 Db 93503 AAGCATTCCTTTTAAAACTGGCAACAGACAGGATGCCCTCTCTCACCACTCTTATTCA 93444
 QY 477 ACATAGTAGTGAAGTTTATGCCAGAGCAATCAGACAGAGAAAGAAATCAAGGCCACC 536
 Db 93443 ACATAGTAGTGAAGTTCTGGCCAGGCGCAATTAGCAGAGAGAAAGAAATTAAGGTAATTC 93384

QY 537 AAATCAATTAAGAGAAAGTCAAACTGTCCCTTTCATGTATGATATGTTATATCTAG 596
 Db 93383 AATTAGAAAGAGAGAAAGTCAAAATGTCCCTGTTTGAGATGAGATGATTTATATCTAG 93324
 QY 597 AAAACCTTAAGACATCAATCCAGAAAGCTCCAGAACTGATATCATTAATTCAGTAAGTTT 656
 Db 93323 AAAACCATATGTCTCAGCCCAAAATCTACTTAAGCTATAGCACTTCAGCAAGTCT 93264
 QY 657 CAGGATCAAACTAATATGACAAATCAGTAGCAGCTATACCAACAGTAGCCAAAG 716
 Db 93263 CAGGATTAATAATCATATGACAAAAATCAACAGCATTCATATACCAAA--TAAGAAA 93208
 QY 717 CTGGAATCAATCAAGAACTCAAACTTTTACAAATGCTGTAA-----AAATATCT 770
 Db 93207 CAGAGACCAAAATCATATGAACTCCCATTCACAAATGTCTTCAAGAGAAATTAATATCC 93148
 QY 771 TAAGATATTCTTCCCAAGAGGTGAGAGACCTCTCAAGAGAAACATCAAAAACAGC 830
 Db 93147 TAGGAATCCAACTTACAAAGGACATGAAAGGCTCTTCAGAGAGAACTCAAAACCATGC 93088
 QY 831 T---GACATCATATGATGACAAACAGTGGAAACATCCATGCTCATGATGGTA 886
 Db 93087 TCAAGAAATTAAGAGAGATCAAAAGAAATGGAAGACTATCCATGCTCATGGTAGGAA 93028
 QY 887 GAATCAATATTGTGAAAAATGACCTATTTGCCAAAAGCAATTTACAAGTTCAATGCAATTC 946
 Db 93027 GAATCAATATCTGTAAAAATGCGCTACTGCGCCAGAGTAATTAATGATTCAGTCCATCC 92968
 QY 947 CCACCAAAATATCATCATCTTCTTCAAGAACTAGAAAAAAACAATTTCTAAATTTCTATA 1006
 Db 92967 CCATCAAGCTACCAATGACTTTCTTCAAGAAATGAAAAAACTACTT-TAAAGTTCTATA 92909
 QY 1007 TGAACACAACCAAAAAAAAAAAAAAAAAAAAAA--CCCGCATAGCCAAAGCAAGCTTAGCAAAA 1066
 Db 92908 TGAACCC-----AAAAAGAGCTGCTACACCAATCATCTTAAGCATTA 92863
 QY 1067 AGAACAAATCTGAGGAGCATCATTTCCATTTGAAACTATATCTACAGAGCTTAATCA 1126
 Db 92862 AGAACAAAGCTGAGGAGCATCAGCTACTGACTTCAAACTATATCTACAGAGCTACAGTA 92803
 QY 1127 CCAAAACATATGAGCAGTCACTAATAAAGTGGCAGCATAGACCAATGAAAAAGAGAGGA 1186
 Db 92802 CCAAAACATATGAGTCTGATCCAAACAGAGATATAGATCAATGAAACAGAACAGAC 92743
 QY 1187 ATCCGAATTAAGCCAAATATATATAGCCAACTGATTTTGAACAAAGCAAAACAAA 1246
 Db 92742 CCTGAGAAATATCTGCGATCTACAACTATCTGATCTTTGAACAACTGAGAAAAACA 92683
 QY 1247 TAAAGTGGGAAAGACATTTAGTTAAACAATGTGCTGAGATTATTTGCAAGCCCAT 1306
 Db 92682 AGCAATGGGAAAGGATTCCTTATTAATTAATGATGCTGGGAAACCTGGCTACCCATAT 92623
 QY 1307 GTGGAATAGAAATCTGATCCCTGCTGCTCACTTAATCAAAAATTTGATACAGAGATG 1366
 Db 92622 GTAGAAAGCTGAATGATGATCCCTT-CTTACACTTATACAAAATTAATTAAGATGG 92564
 QY 1367 ATCAAGACTTAATCTGAGACCTTAACCAATTAATAATTTAGAGATTAACATCAGAAA 1426
 Db 92563 ATTAAAGACTTAATGTTAGACCTTAACCAATTAATAATTTAGAGATTAACATCAGAAA 92504
 QY 1427 ATGCTTTAGACATTTCACTTGAAGCAAGATTTCAATGAGCAAGAACCAAAAGTAATGCA 1486
 Db 92503 ACATTCAGAGCATAGGCAATAGGCAAGATCTTCAATGCTTAACACCAAAAGCATATGCA 92444
 QY 1487 AAAAAACAAAATTAATTAATGAGCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1546
 Db 92443 ACAAAGCCCAAAATTTGACAAATGCAATCTTAATTAATTAATTAATTAATTAATTAATTA 92384
 QY 1547 ACATCTTAGCAGAGCAACAGCAACCAAGCAGTAGAGAAATTTTCAAACTAA 1606
 Db 92383 GAAATCTACATCAGATGAGCAAGCAACCTTAACAAATGAGGAAATTTTTCGAACCTAC 92324
 QY 1607 GCATCTGACTTAAGACTTAATATCCGAATCCACAGAACTCAAAACAAATCAGCAAGAG 1666

OTHER INFORMATION: n = A,T,C or G
 NAME/KEY: misc.feature
 LOCATION: (230157)...(230256)
 OTHER INFORMATION: n = A,T,C or G
 NAME/KEY: misc.feature
 LOCATION: (232299)...(232398)
 OTHER INFORMATION: n = A,T,C or G
 NAME/KEY: misc.feature
 LOCATION: (236552)...(236651)
 OTHER INFORMATION: n = A,T,C or G
 NAME/KEY: misc.feature
 LOCATION: (238789)...(248788)
 OTHER INFORMATION: n = A,T,C or G
 NAME/KEY: exon
 LOCATION: (118288)...(119101)
 OTHER INFORMATION: exon 1C
 NAME/KEY: exon:intron junction
 LOCATION: (151129)...(151130)
 OTHER INFORMATION: exon 5:intron 5
 NAME/KEY: exon:intron junction
 LOCATION: (29248)...(29249)
 OTHER INFORMATION: exon 9:intron 9
 NAME/KEY: exon:intron junction
 LOCATION: (348578)...(348579)
 OTHER INFORMATION: exon 10:intron 10
 NAME/KEY: intron
 LOCATION: (348579)...(381838)
 OTHER INFORMATION: intron 10
 NAME/KEY: intron:exon junction
 LOCATION: (386185)...(386186)
 OTHER INFORMATION: intron 11:exon 12
 US-10-027-983-11

Query Match 12.4%; Score 1068.6; DB 4; Length 392000;
 Best Local Similarity 73.8%; Pred. No. 4.7e-261;
 Matches 1515; Conservative 0; Mismatches 504; Indels 34; Gaps 11;

QY 1 TGTATGAGCCCATGCTTTATATCCAAACAGAGAAAGTATAT--CAAAAAGAA 58
 DB 319240 TTTATGAGGCCAGCATATCTTGATACCAAGCTGCGAGAGCACAAACAAAAGAGA 319239
 QY 59 ACTATAGACAGATACCATGATGATATATATCATGACAGAAATCCCAACAAATATAGCTA 118
 DB 319300 ATTTTNGACCAATATCTCTGATGACATGATGACAAAATCTGAGTAAATATCTGCA 319359
 QY 119 ACCCAATCCACAGCATATCAAGAGATATCCACATGTCAGTGGGTTTATACCA 178
 DB 319360 ACCGAATCCAGAGCATCAAAAAGCTTATCCACATGTCAGTGGGTTTATACCA 319419
 QY 179 GGGTGC-AGGATAGCTTATCATACAGATCAATTAATGATACATCATATAACAGA 237
 DB 319420 GATATCAAGGCTGTTCAACATATGCAAGTCAATTAATCCAGCATATTAACAGA 319479
 QY 228 TTAATAACCAAAATCATCATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAATCT 297
 DB 319480 CCAATGACAAAACCAATGATATCTCATATAGATGCAAAAAGGCTTTGACAAATCT 319539
 QY 298 AACATTTCTTATGATTAATAACCTTCAGCAAAAATGACATAGAAAGACATACCTTAATG 357
 DB 319540 AACCACTTCATGCTTAATAAAGCTCAATATATAGATGATGGAAGCTTATCTGAAA 319599
 QY 358 TTAATAAGCCATATATGACGAGCCACAGCAAACTTATAGTATGAGGAAAAAGTTGA 417
 DB 319600 TTAATAAGACTATCTATGCAAAACCAAGCCAAATTCATATAGATGCAAAAAGCTGG 319659
 QY 418 AAACATTTGCTTGAAGACTGGAACAGACAAAGATG-CTACTTTCACCACTCTATTC 476
 DB 319660 GAGCATTTCCCTTTGAAAACTGGGACAAAGAGGATGCCCTCTCTACCACTCTATTC 319719
 QY 477 ACATAGTAGTGAAGTTTATAGCCAGACATGACAGAAAGAAATCAAGGAGCAACC 536
 DB 319720 ACATAGTGTGGAAGTTTGTGTTATAGGCAATCAGGAGAGAAAGAAATTAAGGATATTC 319779

QY 537 AAATCAATAAAGAGAGATCAAACTGTCCTGTCATGATGATATGATACCTAG 596
 DB 319780 AATTAGAAAAGAGAGATCAAAATGTCCTGTCATGATGATATGATATCTAG 319839
 QY 597 AAAACCTTAAAGACTGATCAGAAAGCTCTAGAACTGATACATTAATTCAGTAAAGTTT 656
 DB 319840 AAAATCCATGCTGTCAGGCAAAATCTCTTAAGCTGATAGCAGCTTCAGCAAGTCT 319899
 QY 657 CAGATACCAATTAATGATACCAAAATGATGACTGCTATACCAAGTACAGTACAG 716
 DB 319900 CAGGATACCAAAATCAATGATGCAAAAATCAAGGCAATCTTATACCAATTAACAGCAA 319959
 QY 717 CTGAGATCAAAATCAAGAACTCAAACTTTATACATAGCT-----GTAAAAAATACCT 771
 DB 319960 CAGAGGCCAAATCATGATGAACTTCCATTCACATGCTTCAAAAGAAATTAATACCT 320019
 QY 772 AAGATATTTCTTACCAAGAGGTGAAAGACCTCTACAAAGAAATCAAAACACAGCT 831
 DB 320020 AAGATCCAACTTACAAAGGATGTAAGGACCTTCAAGAGAACTTCAAAACACAGCTCT 320079
 QY 832 ----GACATCATGATGACAAACAGTGGAAACATCCATGCTCATGATGGGTAG 887
 DB 320080 CAACAAAATTAAGAGATACAAACAAATGCAAGAACATTCATGCTTATGGTGAAG 320139
 QY 888 AATCAATATTGAAAAATGACATATGCGCAAAAGCAATCTCAAGTTCATGCAATTC 947
 DB 320140 AATCAATATCATGAAATATGCGCAATGCTGCGCAAGTATTTATATTCATGCTATCC 320199
 QY 948 CACCAAAATATCATCATCTTCTTCAAGAACTAGAAAAAAACATTTCAATAT 1007
 DB 320200 CATCAAGTACCAATGACTTCTTCAAGAA--TTGAAAAAATCTATTAAAGTTCATAT 320258
 QY 1008 GGAACAACAACCAAAAAAAAAAAAAAAAAAAGCCGCTATGCGCAAGCAAGTACGAAAA 1067
 DB 320259 GGAAC-----CAAAAAAGCCGCTATGCGCAAGTCAATCATTAAGCAAA 320304
 QY 1068 GAACAATCTGAGGACATCACTTACCCATCTTCAACATATCTCAAGGCTATATAC 1127
 DB 320305 GAAACAAGCTGAGGACATCATGCTACTGACTTCAACATATCTCAAGGCTATAC 320364
 QY 1128 CAACAATCATGACCTGACATTAACCTAGGACATAGACCAATGAAAAAGAGAGA 1187
 DB 320365 GAAAAAGCATGCTATGCTTACCAAAAGAGATTAAGCAATGAAACAGAGAGG 320424
 QY 1188 TCCAGAAATTAAGCCCAATTAATTAAGCACTGATTTTGAACAAGCAAAACAAAT 1247
 DB 320425 CTCAGAAATTAAGCCCAATTAATTAAGCACTGATTTTGAACAAGCAAAACAA 320484
 QY 1248 AAAAGTGGGAAAAAGACATCTGATTAACCAATGAGTGTGAGATTTATGGCAAGCCACATG 1307
 DB 320485 GAAATGGGAAAAAGATTCCTTATTTATTAATGATGCTGGGAAAACTGGCTAGCCATATG 320544
 QY 1308 TGGAAATATGAATCTGATCCCTTGTCTCTCACTTAATCAAAAATTTGATACAGATGGA 1367
 DB 320545 TGAAGCTGAATCTGATCCCTTC-CTTACACCTTAATCAAAAATTTATTAATGATGGA 320603
 QY 1368 TCAAGACTTAATCTGAGACTTAACCAATTAATTTAGAGATTAATCATCAGAAAA 1427
 DB 320604 TTAAGACTTAATGTTAGATCAAAAACCATTAATAAACCCTAGAACCAATTA 320663
 QY 1428 TGCTTTAGACATTCCTTATGCAAAAGCTTATATGSCCAAGAACCAAAAGTAAATGCA 1487
 DB 320664 CCAATTAGACATAGGATGAGGCAAGGATTAATGATCTTAATAACCAAAAGCAATGAGCA 320723
 QY 1488 CAAAAACAAAATTAATATAGATGAGCTTAATTAATTAATAAGCTTTTGGCAGCAAA-- 1545
 DB 320724 CAAAACCAAAATTTGCAATATGAGATCTTAATTAATTAAGGCTTTGCAAGCAAAAG 320783
 QY 1546 AACATATCTTATGAGAGCAAAAGACACACCAAGGATGAGAAATTTTCAAACTTA 1605
 DB 320784 AGAACTTACATCAGATGATTAAGAGGCAACCTTACAGAAATGGGAAAAATCTTTGCAATCTTA 320843
 QY 1606 AGCATCTGACTAAGGCTAATATCCGGAATTCACAAAGAACTCAAAACAAATCAGCAAGAA 1665

Db 320844 CTCATTGACAAAGGCGCTAATATCCAGATCTTCAAGAGATCTCAACAAATTTTCAAGAA 320903
Qy 1666 GAAAGCAACATCCATGAAAGAGTGGGCTTAGACATGAATAGACAAATTTCTCAAAA 1725
Db 320904 AAAAACAACACACATCAACAAATGGGCGAAGATATGAACAGACCTTCTCAAAA 320963
Qy 1726 AGATATCAAAATGGCCAAAC-AAAAGAAAAATGCTTAAATCATCTAATGATTTAGGA 1783
Db 320964 AGACATTTATGACACCAACAGACATGAAAAATGTCATCATCTGCGCCATCAGAGA 321023
Qy 1784 AATGTAATCAACATGTAATGATGCTCACTTACTCTGCAAGAAATGTCATATTTA 1843
Db 321024 AATGCAATCAAAACCAATAGATACATCTCAACCACTTAAATGATTTAA 321083
Qy 1844 AAAATCTAAAAATATATATGATGTTGGTGGTCTGTGATTAAGAAACATTTTACATG 1903
Db 321084 GAACTCAGGGGAAAAACGGTGTCTGAGAGATGTGGAATATAGAAAGACTTTTACATG 321143
Qy 1904 CTGGTGGGAATGTAACTTGGGCAACCATATGAGAAAAAGTGTGAAATTTCTTAAAGA 1963
Db 321144 TTGGTGGAGCTGTAACATGTTCAACCATGTGGAATTCAGTGTGGGATTTCTCAGGGA 321203
Qy 1964 ACTAAAGTAGATGACCATTTGATCCAGCAATCCCATTAATATATATATATA 2023
Db 321204 TCTGAACTGAAATACCATTTGACCAAGCATCCCATTT-ACGGGATATATACCAAGG 321262
Qy 2024 TTTATATACCATG 2036
Db 321263 ATTATAAATCATG 321275

RESULT 10
US-10-071-411A-63/C
; Sequence 63, Application US/10071411A
; Patent No. 6797475
; GENERAL INFORMATION:
; APPLICANT: Glenn Barnes
; APPLICANT: Joanne Meyer
; TITLE OF INVENTION: Detection of Polymorphisms in the Human
; FILE REFERENCE: MRI-021
; CURRENT APPLICATION NUMBER: US/10/071,411A
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: 60/267,515
; PRIOR FILING DATE: 2001-02-08
; PRIOR APPLICATION NUMBER: 60/314,248
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 63
; LENGTH: 168174
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(168174)
; OTHER INFORMATION: n = A,T,C or G
US-10-071-411A-63

Query Match 12.3%; Score 1064.6; DB 4; Length 168174;
Best Local Similarity 74.0%; Pred. No. 2.9e-260;
Matches 1519; Conservative 0; Mismatches 499; Indels 35; Gaps 12;
Qy 1 TGTATGAGCAATGTGCTTTATATACCAAAACAGG--AAAGATATACAAAAAGAAA 58
Db 122548 TTTATGAGCGCAGATCATCTGATACCAAAAGCCTGGCAGATACACACAAAAAACA 122489
Qy 59 ACTATAGACCAATGACATGATATATACATGCAAGAAATCCCAACAAATATCTAGCTA 118
Db 122488 ATTTTAAGCAATATCCCTGATGAACATGCAAAAAATCTCAATTAATATCTGCGAA 122429
Qy 119 ACCAATCCAAACGATATCAAGAAATATCCACATGTGCAAGTGGTTTCATACAG 178

Db 122428 ACCGAATCCAGACACATCAAAAAGCTTATCCACGAGATCAAGTGGCATCATCTTG 122369
Qy 179 GGGTGC-AGGATAGGTTTAAATACACAAAGTCAATTAATGTGATCATACATATAACAGA 237
Db 122368 GGATGCAAGGTGTTCAAGTTACAAATCAATTAATGTATATACATATATAACAGA 122309
Qy 238 TTTAAAAAATAATCAATGATCATCTCAATATAGTGTGAAAAAGCATTTGACAAATCT 297
Db 122308 CCAAGACAAAAACCATATGTTATCTCAATATAGTGCAGAAAAAGCCTTTGACAAAAATTC 122249
Qy 298 AACATTTCTTTATATTAATTAACCTTCAGCAAAATCCGACATAGAAAGCATACCTTATG 357
Db 122248 AACCAATTCATGCTAAAACTCTCAATTAATTAATGATTTGATGAGACGATATCAAAA 122189
Qy 358 TAATTAAGCATATATATAGAGGACCCACAGCAAACTTATATCTGAATGGGAAAAAGTTGA 417
Db 122188 TAATTAAGCATATTTATAGCAAAACCCACAAATATCATATGATGAGGCAAAACCTG 122129
Qy 418 AACATTTGCTCTGAGAACTGGAACAAAGCAAGAGATG-CTACTTCAACCACTTCTATTCA 476
Db 122128 AAGCATTCCTTTGAAAACTGGCAGAAAGAGGATCCCTCTCACACTCTCTATTCA 122069
Qy 477 ACATAGTAGTGAAGTTTATGCGCAGAGATCCAGACAAAGAAAGAAATCAAGGGCAACC 536
Db 122068 ACATAGTGTGAAGTTCTGGCCAGGCAATCCAGGAGGAAAGAAATTAAGGGATATT 122009
Qy 537 AATCAATTAAGAGAGAGTCAAACTGCTCCTGTTCACTGATGATGATTTATATCTAG 596
Db 122008 AATTAGGAAAAAGAGAGTCAAAATGTCCTGTTTCAAGATGACATGATTTATATTAG 121949
Qy 597 AAAACCTTAAGACTCATCCAGAAAGTCC-TAGACTGATATCATTAATTCAGTAAAGTT 655
Db 121948 AAAACCCCATTTGCTCAGCCCAAAATCTCCTTAAAGCTGATAGCACTTCAAGAAAGTC 121889
Qy 656 TCGGATACCAACCTAATGTACCAATCATGATGACATGCTATACCAACAGGACCA 715
Db 121888 TCAGGATACCAAAATCAATGTGCAAGATCAAGCAATCTTTATACCAATTAACAGCAA 121829
Qy 716 GCTGAGATCAAAATCAAGAACTCAAACTTTTCAATAGCTGTAAA-----AAATAC 769
Db 121828 ACAGAGGCCAAATCAGAGAGTGAAGTCCATTCATATGTTCAAAAGAAATTAATATAC 121769
Qy 770 TTAAGATATTTCTTACCAGAGAGTGAAGACTCTTACAGAGAAATCAAAACACAG 829
Db 121768 CTAAGATCCAACTTACAAAGGATGTGAAGAGCTCTTCAAGAGAACTCAAAATCACTG 121709
Qy 830 CT-----GACATCATAGTGAACAAAGATGAAACAAATCCCATGCTCATGTAGTGGT 885
Db 121708 CTCAGCAAAATTAAGAGAGACAAACAAATGGAAGAAATCCCATGCTCATGTAGTGA 121649
Qy 886 AGAATCAATATTGGAATAATGACCAATTTGCCAAAGCAATCTTCAATGCAATT 945
Db 121648 AGATGAATATCATTAATAATGGCATACT-CCAGAGTAAATTTATGATTCATATGTCATC 121590
Qy 946 CCCACCAAAATATCATCATCTTCTTCAAGAACTAGAAAAAAACAATCTTAAATTCAT 1005
Db 121589 CCATCAAACTCAACATGACTTTCTTCTCAGAA-TTGAAAAATCTCAGTTAAAGTTCAAT 121531
Qy 1006 ATGGAACAAACCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1065
Db 121530 ATGGAAC-----CAAAAAAGGCCCGCATTCGCAAGAAATCTTAAAGCCAA 121485
Qy 1066 AAGAACAAATCTGAGGACATCATCTTCAATCTTCAATCTTCAAGAGCTTAAATC 1125
Db 121484 AAGAACAAAGCTGAGGACATCATCTTCAATCTTCAATCTTCAAGAGCTTAAATC 121425
Qy 1126 ACCAAAAATCATGCACTGACATTAATTAAGGACATGACCAATGGAAGAAAGAGAG 1185
Db 121424 ACCAAAAATCATGCTGATCTGATCAAAACAGAGATATAGACCAATGGAAGAGAGAG 121365
Qy 1186 AATCCGAATTAAGCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1245

Db 121364 CCTCAGAAATGACACACACATCTACACATCTGATCTTTGACAAACCTGACAAAC 121305
Qy 1246 ATAAAGTGGGAAAAGACATTTCTAGTTACAAATGGTGTGATTAATTTGGCAAGCACA 1305
Db 121304 AAGCATGGGAAAGATTTCCCTATTAAATGATGCTGGGAAAACCTGGCTAGCCATA 121245
Qy 1306 TGTGGAAGATGAACTGGATCCCTGTCTCTCACTTAAATACAAATTTGATACAAGT 1365
Db 121244 GGTAGACAGCTGAAGTGTATCCCTTC-CTTACACCTTTATACAAATTTAATTCAGG 121186
Qy 1366 GATCAAGACTTAAATCTGAGACTTAAACCATTAATTTCTTGAAGATTAATCAAGAA 1425
Db 121185 GATTAAAGACTTAAATGTTAGACTTAAACCATTAACCATTAAGAGAAACCTTAGGCAA 121126
Qy 1426 AATGTTCTAGACATTCATCTTGGGCAAGACTTCAAGGCCAAGAACCCAAATTAATG 1485
Db 121125 TACCATTCAGACATTAAGATGGGCAAGACTTCAAGTCTTAAACACCAAAACCAATG 121066
Qy 1486 AACAAACCAAAATTAATAGATAGAGCTTAAATTAATTAATTAATTAATTAATTAAT 1545
Db 121065 AACAAAGCCAAATAGACAAATGGATCTTAATTAATTAATTAATTAATTAATTAAT 121006
Qy 1546 AACAACTATTAGACAGACAAACACACACACCTGAGTGAAGAAATTTTCAAACTA 1605
Db 121005 AGAACTACATCAGAGTGAACAGGCACTTACAGATGGGAAATTTTCAATCTA 120946
Qy 1606 AGCATCTAGACTAAGACTAATCCGGAATCCCAAGAACTTAAATTAATTAATTAATTA 1665
Db 120945 CTCACTGCAAAAGGCTAATTCAGAAATCTTAAACCAATTAATTAATTAATTAAT 120886
Qy 1666 GAAAGCAAACTCCATGAAAGAGTGGCTAAGACATGAATTAATTAATTAATTAATTA 1725
Db 120885 AAAACCAAACTCCATGAAAGAGTGGCTAAGACATGAATTAATTAATTAATTAAT 120826
Qy 1726 AGATATCAATAGCCCAAC-AAAACGAAAAATCTTAATCACTAATTAATTAATTA 1783
Db 120825 AGACCTTTATGACAGCAACAGACACATGAATAATCTCATCTGAGCCATCAGAGA 120766
Qy 1784 AATGTAATCAACATGTAATGATGATACACCTTACTCTGCAAGAAATGCTAATTAAT 1843
Db 120765 AATGCAATCAAAACAT 120706
Qy 1844 AAAATCTAAAAATTAATAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1903
Db 120705 AAAAGTCAAGAAACACAGGTGTGAGAGAGGTGTGAGAAATGGAACCTTTTAAT 120646
Qy 1904 CTGGTGGGAATGTAATCTGGCAACCATATGAGAAACAGTGTGAAATTTCTTAAGA 1963
Db 120645 TTGGTGGGACTGTAACTGTTCACCTTGTGGAAGACACTGTGTGCAATTTCTCAG 120586
Qy 1964 ACTAAAGATAGATCGACATTTGATCGAATCCCATTAATATATATATATATATAT 2023
Db 120585 TCTAAGACTAAGAAATACATTTGACCCAGCATCCCAT-ACTGGTATATATCCCAA 120527
Qy 2024 TTTATATACCATG 2036
Db 120526 ATTATATATCATG 120514

RESULT 11
US-10-071-411A-2/c
; Sequence 2, Application US/10071411A
; Patent No. 6797475
; GENERAL INFORMATION:
; APPLICANT: Joanne Meyer
; TITLE OF INVENTION: Detection of Polymorphisms in the Human
; FILE REFERENCE: MRI-021
; CURRENT APPLICATION NUMBER: US/10/071,411A
; PRIOR APPLICATION NUMBER: 2002-02-07
; PRIOR FILING DATE: 2001-02-08

; PRIOR APPLICATION NUMBER: 60/314,248
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 168273
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(168273)
; OTHER INFORMATION: n = A,T,C or G
US-10-071-411A-2

Query Match 12.3%; Score 1064.6; DB 4; Length 168273;
Best Local Similarity 74.0%; Pred. No. 2.9e-260;
Matches 1519; Conservative 0; Mismatches 499; Indels 35; Gaps 12.

Qy 1 TGTATGAAAGCAATGCTACTTTAATACCAAAACAGG--AAAGATATACAAAAAGAA 58
Db 122647 TTTATGAGGCCAGCATGATCCCTGATACCAAGCTGGCAGATACACACAAACAA 122588
Qy 59 ACTATGACAGTACCATGATGAATATACATGCAAAATCCCAACAAATACTAGCTA 118
Db 122587 ATTTTGAACCAATTCCTGATGAACATGATGCAAAATCTCTCAATAAATACTGG 122528
Qy 119 ACCCATCAACAGCATATACAAAGATATCCATTCGTAAGTGGTTTCAATACAG 178
Db 122527 ACCGAATCAGACGACATCAAAAGCTTATCCAGAGATCAAGTGGCATCATCTG 122468
Qy 179 GGTGCG-AGATATAGTTTAAATATACACAAAGTCAATTAATGATATACATCAATTA 237
Db 122467 GATGGAAGCTGTGTTCAATACATACCAATATATATATATATATATATATAT 122408
Qy 238 TTTAAACCAAAATCATGATCATCTCATATGATGCTGAAAAGCAATTTGACAAATCT 297
Db 122407 CCAAGACAAACCAACGATTAATCTCAATGATGCAAGAAAGCCCTTGAACAAATTC 122348
Qy 298 AACATTTCTTTATGATTAATTAACCTTACAGCAAAATGACATATGAAGGACATPCT 357
Db 122347 AACACCATTCATGCTAATAATCTCTCATATTAATTAATGATGATGAGAGCTTCA 122288
Qy 358 TATATAACCATATATATGACGACACCAACCAATTAATCTGAATGGGAAAGTTGA 417
Db 122287 TATATAGACTTTATATGACAAACCAACCAATATCATCTGAATGGGAAACCTG 122228
Qy 418 AAACATTTGCTGAGAACTGGAACCAAGCAAGATG-CTACTTTCACCACTTATATCA 476
Db 122227 AAGCATTCCTTTGAAATCTGCAACAGGAGTGGCTCTCTCACCACTCTATTTCA 122168
Qy 477 ACATAGTAGTGAAGTTTATAGCCAGAGCAATGACACAGAAAGAAATCAAGGCAACC 536
Db 122167 ACATAGTGTGGAAGTTCTGGCCAGAGCAATAGGCAAGAAAGAAATTAAGGTAAT 122108
Qy 537 AAATCAATTAAGAGAGAGTCAACTGCTCCCTGTTCACTGATGATATGATTTATACCT 596
Db 122107 AATTGGAAGAGAGAGTCAATGCTCCCTGTTTCCAGATGACATGATTTATTTAG 122048
Qy 597 AAAACCTTAAAGTCTATCAGAAAGCTCC-TGAACTGATATCAATTAATTCAGTAAAT 655
Db 122047 AAAACCCATTTGCTCAGCCCAAAATCTCTTAAAGCTGATGGCACTTCAGCAAT 121988
Qy 656 TCAGATACAAACTAATGATACAAATTCAGTACAGTCTATACCAACAGTGAACCA 715
Db 121987 TCAGATACAAATCAATGATGAGAAATTCAGAAATTTCTTATACCAATTAACAG 121928
Qy 716 GCTGGAATCAATCAAGAACTCAAACTTTTCAATAGCTGTAAA-----AAAATAC 769
Db 121927 ACAGAGAGCAATTCAGAGTGAATCTCCATTTCAATTTGCTACAAAGAAATTAAT 121868
Qy 770 TTAAGAAATTTCTTACCAAGAGAGTGAAGAGCTCTTCAAGAGAAACCTCAACAAAC 829
Db 121867 CTAGAAATCAACTTACAAAGGAGTGAAGAGCTTTCAAGAGAACTACAAATCACTG 121808

QY 830 CT-----GACATCATAGTACGAAAGAGTGAAGAACATCCCATGCTCATGATGGT 885
 DB 121807 CTCAACCAAAATTAAGGAAGACAAACAAATGGAAGAACATCCCATGCTCATGATGAGA 121748
 QY 886 AGAATCAATATTGGAATGACATATTGCGCAAGCAATCTCAAGTTCAATGCAATT 945
 DB 121747 AGAATCAATATTGGAATGACATATTGCGCAAGCAATCTCAAGTTCAATGCAATT 121689
 QY 946 CCCACCAAAATATCATCATCTATCTTCAAGAACATGAAAAAAACAATTTCTAAATTCAT 1005
 DB 121688 CCATCAAACTACCAATGACCTTTCTTCTGAGA--TTGGAAAAATCTACGTTAAAGTTTAT 121630
 QY 1006 ATGGAACCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1065
 DB 121629 ATGGAAC-----CAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 121584
 QY 1066 AGAATCAATATCTGAGGATCATCTTCCCATCTTCAAACTATCTCAAGGCTATATC 1125
 DB 121583 AGAATCAATATCTGAGGATCATCTTCCCATCTTCAAACTATCTCAAGGCTATATC 121524
 QY 1126 ACCAAACATCATGACATGACATGAACTAGGACATAGACCAATGGAAGAGAG 1185
 DB 121523 ACCAAACATCATGACATGACATGAACTAGGACATAGACCAATGGAAGAGAG 121464
 QY 1186 AATCGAATTAAGCCCAATTAATTATGCACTGATTTTGAACAAAGCAACCAAAAC 1245
 DB 121463 CCCTCAACAAATGACACACACATCTACCAACCATCTGATCTTTGACAAACCTGACAAAC 121404
 QY 1246 ATAAAGTGGGGAAGAAAGCATTTCTAGTTAACAATGAGTGGAGTTATGCGAAGCCACA 1305
 DB 121403 AAGCAATGGGGAAGAAAGCATTTCTAGTTAACAATGAGTGGAGTTATGCGAAGCCACA 121344
 QY 1306 TGTGGAAGATGAACTGATCCCTTGTCTCTCACTTAATCAAAAAATGATACAAAGT 1365
 DB 121343 GGTAGACAGCTGAAGCTGATCCCTTGTCTCTCACTTAATCAAAAAATGATACAAAGT 121285
 QY 1366 GATCAAGACTTAATCTGAGACCTTAAACCAATTAATTTAGAGATTAACATCAGAA 1425
 DB 121284 GATTAAGACTTAATCTGAGACCTTAAACCAATTAATTTAGAGATTAACATCAGAA 121225
 QY 1426 AATGCTTCTGACATCTTCACTTAAGCAAGATTTGATGCGCAAGAACCAAAAGTAAATGC 1485
 DB 121224 TACCATTCAGACATTAAGCAAGATTTGATGCGCAAGAACCAAAAGTAAATGC 121165
 QY 1486 AACCAAAACCAAAATTAATAGATAGACTTAATTAATTAATTAATTAATTAATTAAT 1545
 DB 121164 AACCAAAACCAAAATTAATAGATAGACTTAATTAATTAATTAATTAATTAATTAAT 121105
 QY 1546 AACCAATCAATGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1605
 DB 121104 AGAAATCAATGACAGTGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCTA 121045
 QY 1606 AGCATCTGACATGACATTAATATCCGAATTCAGAGAACTCAACCAAAATCAGCAAGAA 1665
 DB 121044 CTCAATCTGACAGAGGCTTAATATCCGAATTCAGAGAACTCAACCAAAATTTAAGAA 120985
 QY 1666 GAAAGCAAAACATCTGAGAGAGTGGGCTTAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1725
 DB 120984 GAAAGCAAAACATCTGAGAGAGTGGGCTTAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 120925
 QY 1726 AGATATCAAAATGGCCAA--AAACAGAAAAAAATGCTTAACATCACTAATGATTAGGA 1783
 DB 120924 AGAGCTTATGACAGCCAAACAGACATGAAAAAAATGCTCATCATCTGCGCATCAGGA 120865
 QY 1784 AATGTAATCAACCTGTAATGCAATCACTTAATCTTCCGAAGATGCTCAATTAATTA 1843
 DB 120864 AATGTAATCAACCTGTAATGCAATCACTTAATCTTCCGAAGATGCTCAATTAATTA 120805
 QY 1844 AAAATCTAAATTAATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1903
 DB 120804 AAAATCTAAATTAATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 120745

QY 1904 CTGTTGGGAATGTAATCTTGCGCAACCACTATGGAAGAAACAGTGTGAATTTCTTAAGA 1963
 DB 120744 TTGTTGGGAGCTGTAACTAGTTCAACCTTTGTGGAAGACACTGGGAATTCACAGGA 120685
 QY 1964 ACTAAAGTATGATGACCATTTGATCCAGCAATCCCAATTAATTAATTAATTAATTAAT 2023
 DB 120684 TCTGAACCTGAATTAACATTTGACCGGATCCATTTACTGGTATATACCCAAAG 120626
 QY 2024 TTATATACCATG 2036
 DB 120625 ATTATATCATG 120613

RESULT 12
 US-09-816-095-3/c
 ; Sequence 3, Application us/09816095
 ; Patent No. 6664084
 ; GENERAL INFORMATION:
 ; APPLICANT: GAN, Weiniu
 ; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES.
 ; FILE REFERENCE: C1001147
 ; CURRENT FILING DATE: 2001-03-26
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 99916
 ; TYPE: DNA
 ; ORGANISM: Human
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(99916)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-816-095-3

Query Match 12.3%; Score 1064.2; DB 4; Length 99916;
 Best Local Similarity 74.0%; Pred. No. 2.6e-260;
 Matches 1493; Conservative 0; Mismatches 488; Indels 36; Gaps 10;

QY 1 TGTATGAAGCAATGTCATTTAATACCAAAACCGAAGAAAGATTA--CAAAAAAGAA 58
 DB 21190 TTTATGAGGCGACGATCATCTCTGATACCAAGCGGACAGACACAGCCAAAAAGAGA 21131
 QY 59 ACTATGACAGTACACATGATGATATATACATGACAGAAATCCCAACAAATACTAGCTA 118
 DB 21130 ATTATGACCAATATCTTGTATGACATGATGACAAAAATCTCAATTAATTAATCTGCA 21071
 QY 119 ACCCAATCAACAGCATATCAGAGAAATATCCACATTTGCAAGTGGGTTTCATPACAG 178
 DB 21070 ACCCAATCCAGCAGCATCAACAAAGCTTATTCACATGATCAAGTGGGCTTCATCCCTG 21011
 QY 179 GGGTGC--AGATAGGTATTAATACATACAGCAAGTCAATTAATGATATCATCACTAACAAG 237
 DB 21010 GGAATGCAAGGCTGTTCAATATATGCAATCAATTAATGATATCACTAATTAACAAGAA 20951
 QY 238 TTAATAACAAAAATCACATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAATCT 297
 DB 20950 CCAAGCAAAAAACCAATGATTTATCTCAATAGATGCAAGAAAGGCTTTGACAAATTC 20891
 QY 298 AACATTTCTTATGATTAATAAATCTTGACAAATGACATGAAAGACATCACTTAATG 357
 DB 20890 AACCAAGCTTCATCTTAATAAATCTCAATTAATTAAGTATGATGAGGACGTATCTCAAAA 20831
 QY 358 TAATTAAGGCTATATGAGAGACCCACAGCAACATTAATTAATGAAATGGGAAAAAGTTGA 417
 DB 20830 TAATTAAGGCTATATGAGAGACCCACAGCAATCAATATCAATGAAATGGGAAAAAGTTG 20771
 QY 418 AACATTTGCTCTGAGAGACTGAAACAGCAAGAGATG--CTACTTTCACCACTTATTTCA 476
 DB 20770 AGGATTTCCCTTTGAAAAACGGGACAGAGACAGGATGCCCTCTCTACACACTCTATTTCA 20711

QY	4777	ACATGTAGTGGAAAGTTTGTTCGACGACCAATCAGACAAGAAAGAAATCAAGGGCACCC	536
Db	20710	ACATGTGTGGAAAGTCTTGCCACGGGCAATTTAGCGAAGAAAGCACTAAAGGGTATTC	20651
QY	537	AAATCAATTAAGAGGAAGTCAAACTGTGCCCTGTTCACTGATGATATGATGTATCTAG	596
Db	20650	AATTAGGAAAAAGAGGGAGTCAAATTTGTCCGTGTTTGACAGACGACATGATTTGTATCTAC	20591
QY	597	AAAACCTTAAGAGCTCATCCGAAAGCTCCTGAACTGATACATTAATTCACTTAAGTTT	656
Db	20590	AAAACCCCAATGTCTCAGCCAAAATCTCCTTAAGCTGATTAAGCAACTTCAGCAAGCTC	20531
QY	657	CAGGATACAAACTAAATGTACACAATTCAGTACGCTCTTATACCAACGTCACAG	716
Db	20530	CAGGATACAAAATCAATGTACAAAATTCACAGCACTTTATACCAATTAACAGACAA	20471
QY	717	CTGAGATCAAAATCAAGAACTCAACACTTTTAAATAGCT-----GTAAAAAATCT	770
Db	20470	CAGAGAGCCAAATCATGAGTGAATCTCCATGTACAACTTGCTTCAAAAGTAAATTAATCT	20411
QY	771	TAAAGATATTCTTACCCAGAGGTGAAGACTCTTCAAGAGAAACTACAAACACAGC	830
Db	20410	TAGGAATCCAACTTACAAAGGAGCGTGAAGGACCTCTTCAAGAGAACTACAAACACCTGC	20351
QY	831	-----TGACATCATAGATGACACAAACAAGTGGAAAAACAATCCATGCTCATGGATGGGTA	886
Db	20350	TCAATTAATAAATAAGAGATACAAACAATAGAAAGAAATTCCTCATGCTCATGGTGAAGAA	20291
QY	887	GAATCAATATTGTGAATAATGACATATTGTCGAAAAAGCAATCTACAAAGTTCAATGCATTC	946
Db	20290	GAATCAATATTGTGAATAATGGCCATACGCGCCAAAGTAAATTTATGATTCATATGCATCC	20231
QY	947	CCACCAAAATATCATCATCTTCTTCCACAGACTAGAAAAAAACAATTCMAATTTCATA	1006
Db	20230	CCATCAAGCTACCAATGACTTCTTCCACAGAA--TTGAAAAAAACTACTTAAAGTTTATA	20172
QY	1007	TGGAAACAACACCAAAAAAABAAAAAACCCTGATGCCAAAGCAGACCTTACCAAA	1066
Db	20171	TGGAAAC-----CATAAAAAGGCCACATGTGTAAGTCAATCTCTAACCCAAA	20126
QY	1067	AGAACCAATCTGAGGCGATCACAATTACCCTTCAACTATATCTATACAAAGGCTATATCA	1126
Db	20125	AGAACCAAGCTGAGGCGATCACACTTACTGACTTCAACTATATCTATACAAAGGCTATACGTA	20066
QY	1127	CCAAAAATCATATGCGCACTGACATTAACCTAGGCACTATGACCAATGAAAGAAAGAGAGA	1186
Db	20065	CCAAAAACAGACTGTACTGTGACCAAAACAGAGATATGATCAATGGAACAGAACAGAGC	20006
QY	1187	ATCCGAAATTAAGCCAAATTAATTAATATAGCCACTGATTTTGAACAAGCAACAAAAACA	1246
Db	20005	CCTGCAAAATTAAGCTGCGCATATCTACAACTATCTGATCTTGAACAACCTGAGAAAAACA	19946
QY	1247	TAAAGTGGGGAAGAAGACATTCTAGTTTAACAATAGGTGCTGAGATAATTGGCAAGCCACT	1306
Db	19945	AGCAATGGGGAAGAAGATTCCCTTAATTTAATTAATATGTGCTGGAAAACTGGCTAGCCATAT	19886
QY	1307	GTGGAAGAAATGAAACTGGATCCCTGTCTCTCACTTAATACAAAAATTTGATCAAGATGG	1366
Db	19885	GTAGAAAGCTGAAGACTGATCCCTT--CCTTACACTTATACAAAAATTAATCAAGATGG	19827
QY	1367	ATCAAAAGCTTAATCTGAGACCTTAAACCACTAAAAATTTCTGAAAGATTAACTCAGAAA	1426
Db	19826	-----AAACTTAAACGTTAGACCTTAAACCACTAAAAACCTTGAAGAAAACTTAGGCAT	19772
QY	1427	ATGCTCTAGACACTTCACTTAGGCAAAAGACTCATGGCCAGAAACCCAAAGCTAATATGCA	1486
Db	19771	ACCATTCAGACACTAGGCAATGGGCAAGGACTTCATGCTTAAACACCAAAACATTTGGCA	19712
QY	1487	ACAAAAACAATAATTAATAGTATGACTTAATTAACCTTAAAGAGCTTTTGCACACAAA	1546
Db	19711	ACAAAAGCCAAATTTGACAAATGGGATCTTAATTAACCTTAAGAGCTTTCTGCACACACAAA	19652
QY	1547	ACATATATTGACAGACAAACAGACCAACCGAGTGAAGAAAACTTCCACAACTTA	1606

Db	1951	GAAACCTACATCAGAGTGAACAGGCACCTTCAAAATGGAGAAATATTTGCACTAC	19592
Qy	1607	GCATTCGACTAAGACTAAATATCCGGAATCCACAAGAACTCAACAAATCAGCAGAAG	1666
Db	19591	TCATCTGACAAAGGGCTAATATCCAGAAATCTACATGAATCAAACTTTCACAGAAA	19532
Qy	1667	AAAGCAAAACATTCCTCCATGGAAGAAGGGGCTAAGGACATGATATAGCAATTCCTCAAAAGAA	1726
Db	19531	AAACCAACCAACCCCATCAACAAGAGGGGGAAGGACATGAATAGACACTTCTCAAAAGAA	19472
Qy	1727	GATATCAAAATGGCCCAACAACAAGGAAAAATATGCTTAACATCACTAAATGAATAGAGGAAT	1786
Db	19471	GACATTTATGCAAGCCAAAAGACATATGAAAAAAC-GCTCATCATCTGGCCATCAGAGAAAT	19413
Qy	1787	GTAATCAACACTGTAAATGCGATACCACTTATCTCTGCAAGAAATGGTCATATTTAAAA	1846
Db	19412	GCAAAATCAAAACCAAGATAGATATCCATCTCAACAGATTTAGAAATGGATCATTTAAAA	19353
Qy	1847	ATCTAAAAATATATAGATGTTTGGTGGTCTGTGCGATATAAGAACACTTTTACCTGCTG	1906
Db	19352	GTCAGGAAACACACAGGTGCTGGAAGAGATGTGAGAAATAGAAACACTTTTACCTGTGG	19293
Qy	1907	GTCGGAATGTAAACTTGGCCCAACAACATAAGAAAAACATGTGAAAAATTTCTTAAGAACT	1966
Db	19292	GTGGGACTGTAACTATGTAAAAACATGTGGAAATCAATGTGGGGCATTTCTCAGAGAACT	19233
Qy	1967	AAAAGTAGATCGACCATTTGATCCAGCAATCCATTA	2003
Db	19232	AGAACTAGAAATATCATTTGATTCAGCCATCCATTA	19196

RESULT 13

US-10-071-411A-63

Sequence 63, Application US/10071411A

Patent No. 6797475

GENERAL INFORMATION:

APPLICANT: Glenn Barnes

APPLICANT: Joanne Meyer

TITLE OF INVENTION: Detection of Polymorphisms in the Human

TITLE OF INVENTION: 5-Lipoxygenase Gene

FILE REFERENCE: MRI-021

CURRENT APPLICATION NUMBER: US/10/071,411A

CURRENT FILING DATE: 2002-02-07

PRIOR APPLICATION NUMBER: 60/267,515

PRIOR FILING DATE: 2001-02-08

PRIOR APPLICATION NUMBER: 60/314,248

PRIOR FILING DATE: 2001-08-21

NUMBER OF SEQ. ID NOS: 66

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 63

LENGTH: 168174

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)...(168174)

OTHER INFORMATION: n = A,T,C or G

US-10-071-411A-63

Query Match	12.3%	Score 1058.6;	DB 4;	Length 168174;
Best Local Similarity	73.4%	Pred. No. 9.8e-259;		
Matches 1526;	Conservative 0;	Mismatches 514;	Indels 39;	Gaps 12;

1 TGTATGAAGCCTATGTCACCTTAAATACCAAAACGAGAAAGATATA--CAAAAAGAAA

159587 TTTATGAGGCCACATCATCTCTGATATCCAAAGCCGGGCGAGAGACAAAACCAAAAAGAGG

59 ACTATAGACCAAGTACCACTGATGAATATATACATGCAAGAAATCCCAACAAATATCTAGCTA

159647 ATTTTAGACCAATATCTCTTGAAGGAACATTTGATGCAAAATCTCTCAATTAATACTGGCAA

119 ACCCAATCCAAAGCATATCAAGAAAGATATATCCATTCCTCAAGTGGGTTTCATACCAG

Db 159707 ACCAAATCCAGACGACATCAAAAAGCTTATCCACATGATCAAGTGGCTTCACTCCCTA 159766
 Qy 179 GGGTGC -AGGATAGGTTAAACATACACAAAGTCATTAATGATGATCATCACTATAACAGAA 237
 Db 159767 GGATGCAAGGCTTGTCAACATATGCAAAATCAATTAATGTAATCCAGATATTAACGAA 159826
 Qy 238 TTTAAAAAATAATCATGATCATCTCATATGATGCTGAAAAAGCATTTGACAAATCT 297
 Db 159827 CCAAAAGCAAAAACACATGATTTATCTCAATGATGAGAAAAGCCCTTTGACAAATTC 159886
 Qy 298 AACATTTCTTATGATTTAAACCTTCAGAAAATGACATGAAAAGACATACCTTAATG 357
 Db 159887 AACAACTTCATATGTTAAAACTCTCAATTAATTAATGATTTGCTGGACGATATCAAAA 159946
 Qy 358 TAAATAAAGCATATATGACGACCCACAGAAACATTAATAGTAAAGGAAAAAGTTGA 417
 Db 159947 TAAATAGAGCTATATATGTCAAACCCACAGCAATATCATATGGAATGGCAAAAACCTG 160006
 Qy 418 AACATTTGCTTGAAGACTGGAACAAAGCAAGAGATG -CTACTTTCACCACTTCTATTCA 476
 Db 160007 AAGATTTCCCTTTGAAAATTTGGCAACAAGACAGGATGCTCTCAGCACTCCATTTCA 160066
 Qy 477 ACATAGTAGGAGGATTTAGCCAGAGCAATCAGAAAGAAAGAAATCAAGGCAACC 536
 Db 160067 ACATAGTCTGGAAGATCTGTCAGGCAATCAGGAGAAAGAAATTAAGGGATATTC 160126
 Qy 537 AATCAATTAAGAGAGAGTCAAACTGCTCCCTGTTCACTGATGATGATGTTGATATCTAG 596
 Db 160127 AATTAGAAAAAGAGAGTCAAAATTTGCTGTTTGAAGATGATGATTTTATATCTAG 160186
 Qy 597 AAAACCTTAAGACTCATCCAGAAAGCTCTTGAAGCTGATATCAATTAATTCAGTAAAGTT 656
 Db 160187 AAAACCTTATGCTCAGCCCAAAACCTCTTAAGATGATGAGCACTTCAGCAAGTCT 160246
 Qy 657 CAGGATCAAACTTAATGATCAAACTAGTACAGTCTATACCAAGCAAGTACCAAG 716
 Db 160247 CAGGATCAAAATCAATGATCAAAATTCACAAAGATTTTATACCAACCAAGCAAA 160306
 Qy 717 CTGAGATCAAAATCAAAAGTCAAACTTTTACATATGCTGTAAA-----AAATATCT 770
 Db 160307 CAGAGACCAATATATGATGAACTCCCATTCACAAATGCTTCAAGAGATTAATATCC 160366
 Qy 771 TAAAGATATTTCTTACCAAGAGGATGAAAGACTCTTCAAGAGAAATCTACAAACACAGC 830
 Db 160367 TAGGAATCCACCTTACAAAGGATGTAAGCACTCTTCAAGAGAACTACAAACCACTGC 160426
 Qy 831 T-----GACATCATGATGACCAAAAGTGAAGAACTCCATGCTCATGATGGGTA 886
 Db 160427 TCAACGAATTAAGAGAGATCAAAACAAATGAAAGAACTTCCATGCTCATGGGTAGGAA 160486
 Qy 887 GAATCAATATTTGAAAAATGCAATTTGCCAAAAGCAATCTACAAAGTCAATGCAATTC 946
 Db 160487 GAATCAATATTTGAAAAATGCAATTTGCCAAAAGTCAATTTGCAATGCAATTC 160546
 Qy 947 CCAACCAAAATATCATCATCTTCTTCAAGAACTAGAAAAAAACATTTCTAAAATTCATA 1006
 Db 160547 CCATCAAGCTACCAATTAATCTTCTTCACTAAA -TTGAAAAAACTACTTTAAAGTTCA 160605
 Qy 1007 TGAACACAAACCAAAAAAAGAAAAAAGCCGATAGCCAAAGCAAGCACTTGCAAAA 1066
 Db 160606 TGGAAAC-----CAAAAAGAGCCGCAATCAACCAAGTCAATCTTGAGCCAAA 160651
 Qy 1067 AGAACAATCTGAGGATCATCTTACCATCTTCAAACTATCTACAAAGGCTAATATCA 1126
 Db 160652 AGAACAAGGCTGAGGATCATCTTACCATCTTCAAACTATCTACAAAGGCTAATATCA 160711
 Qy 1127 CCAAAACATCATGCACTGACATTAAGCACTAGCAATAGCAATGAAAAAGAGAGA 1186
 Db 160712 CCAAAACAGCTTGTACTGTTACCAAAAACAGAGATATGATCAATGGAACAGAACAGGC 160771
 Qy 1187 ATCCAGAAATTAACCAATTAATATAGCCAACTGATTTTGCAGAAAGCAAAACAAA 1246

Db 160772 CCTCAGAAATTAATGCTGATATGACAACTGTCTGATCTTTGACAAACCTGAGAAAAACA 160831
 Qy 1247 TAAAGTGGGAAAAAGACATTTCTAGTTAACAAAATGCTGAGATTAATTTGCAAGCCACAT 1306
 Db 160832 AGAAATGGGAAAAAGATTTCCCTATTTATTAATGCTGAGAAAACTGGCTAGCCATAT 160891
 Qy 1307 GTGGAATAATGAAATCTGATTCCTTCTGCTCTGCTTAAATCAAAATTTGATACAAAGATG 1366
 Db 160892 GTGGAATAATGAAATCTGATTCCTTCTGCTCTGCTTAAATCAAAATTTGATACAAAGATG 160950
 Qy 1367 ATCAAGACTTTAATCTGAGACCTTAACCAATAAAATTTCTAGAAATTAATCATGAGAAA 1426
 Db 160951 ATTAAGACTTTAATCTGAGACCTTAACCAATAAAATTTCTAGAAATTTGATACAAAGATG 161010
 Qy 1427 ATGCTTGAACATTTCACTTGAAGCAAGCTTCAATGAGCAAGAACCCAAAGTAATGCA 1486
 Db 161011 ACATTTGAGCAATGAGCAATGAGCAAGCTTCAATGAGCAAGAACCCAAAGTAATGCA 161070
 Qy 1487 AAAAAAATAATTAATTAATGATGAGCTTAATTAATTAATTAATTAATTAATTAATTAATTA 1546
 Db 161071 AAAAAAATAATTAATTAATGATGAGCTTAATTAATTAATTAATTAATTAATTAATTAAT 161130
 Qy 1547 ACATCACTTGAAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1606
 Db 161131 GAATCTACATCAGAGTGAATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 161190
 Qy 1607 GCATCTGATTAAGGATTAATTAATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1666
 Db 161191 TCATCTGACAAAGGCTTAATTAATGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 161250
 Qy 1667 AAGCAAAACATCCCATGAAAGAGTGGGCTTAAGGACATGATGACATTTCTCAAAAGAA 1726
 Db 161251 A-----AAACAAACCCCATCAAAAGTGGGCAAAATTAATGAAAGCACTTCAAAAGAA 161306
 Qy 1727 GATATCAAAATGAGCAAC--AAACAGAAAAAATGCTTAATCATGATTAATGATAGGAA 1784
 Db 161307 GACATTTATGACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 161366
 Qy 1785 ATGTAATCAACACTGTAATGCAATACCACTTACTCTGCAAGAAATGCTCATTAATTTAA 1844
 Db 161367 ATGTAATCAAAACCAAGATGATGATCAATCTGCAAGTAAATGACATCTT--A 161424
 Qy 1845 AATCTTAATAATTAATGATGTTGGTGGTCTGCTGATTAAGAAACATTTTACACTGC 1904
 Db 161425 AAGTCAGAAACCAAGATGTTGAGAGGACATGAGAAATGAGAAATGCTTTTACACTGT 161484
 Qy 1905 TGTGGGAATGTAATCTTGGCAACCACTATGAAAACAGTGGAAATTTCTTAAGGAA 1964
 Db 161485 TGTGGGAGGTAAATTAATTAATGATGATGTTGAAAGCAAGTGGCAATTTCTTAAGGAT 161544
 Qy 1965 CTAAAGTATGATGACCATTTGATCCAGCAATCCCATTAATATGATTAATATATATAT 2024
 Db 161545 CTAAAGTATGAAATTAATGATTTGACCAAGCAATCCCAT--ACTGGTATATACGAAAAA 161603
 Qy 2025 TTATATACATGGAATTAATCACTGACCAATTAATTAATTAATTAATTAATTAATTAAT 2063
 Db 161604 TTACAAATGATGCAATTAATTAAGACATGACACAGAA 161642

RESULT 14
 US-10-071-411A-2
 ; Sequence 2, Application US/10071411A
 ; Patent No. 6797475
 ; GENERAL INFORMATION:
 ; APPLICANT: Glenn Barnes
 ; TITLE OF INVENTION: Detection of Polymorphisms in the Human
 ; FILE REFERENCE: 5-Lipoxigenase Gene
 ; CURRENT APPLICATION NUMBER: US/10/071, 411A
 ; CURRENT FILING DATE: 2002-02-07
 ; PRIOR APPLICATION NUMBER: 60/267,515
 ; PRIOR FILING DATE: 2001-02-08

```

: PRIOR APPLICATION NUMBER: 60/314,248
: PRIOR FILING DATE: 2001-08-21
: NUMBER OF SEQ ID NOS: 66
: SOFTWARE: FastSeq for Windows Version 4.0.
: SEQ ID NO 2
: LENGTH: 168273
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (1)_(168273)
: OTHER INFORMATION: n = A,T,C or G
: JS-10-071-411A-2

```

Query Match	12.3%;	Score 1058.6;	DB 4;	Length 168273;
Best Local Similarity	73.4%;	Pred. No. 9.8e-259;		
Matches 1526;	Conservative	0;	Mismatches 514;	Indels 39; Gaps 12;

Qy	1	1GTAATGAAGCCCATGTGCACTTTAAATCCAAAACCAAGAAAAGATTAT--CAAAAAGAAA	58
Db	159686	TTTATGAGGCCCGACATCATCTTGATATCCAAAGCCGGGACAGACAAAACCAAAAAAGAGG	159745
Qy	59	ACTATGAGCCAGTACCCTGATGATGATATATCATGACAGAAATCCCAACAAATATCTAGCTA	118
Db	159746	ATTTTAGACCAATATCTCTTGAGGAACTGATGCCAAATCTCTAATATTAATCTGGCAA	159805
Qy	119	ACCCAAATCCAAAGCATATCAAGAAATATATCCACTTGTGAATGGGTTTCATACCAAG	178
Db	159806	ACCAAAATCCAGACGACATCAAAAAACCTTATCCATGATGATCAAGTAGGCTTATCTCTA	159865
Qy	179	GGGTGC-AGGATAGGTTAATACATACACAAGTCATTAATGTGATATCATCATAAACAGAA	237
Db	159866	GGATCCAGGCTGTGTTCAACATATGGAATTCATTAATGTAAATCCAGCATATTAACAGAA	159925
Qy	238	TTAAAAAATAATCATCATGATCATCTCAATAGATCCTGAAAAAGCATTTGACAAATCT	297
Db	159926	CCAAAAGACAAAAACCATGATTTATCTCAATAGATCAGAAAAAGGCTTTGACAAATATTC	159985
Qy	298	AACATTTCTTTATGATTTAAACCTTCAGGAAAATGCAATAGAAAGACATACCTTAATG	357
Db	159986	AACAACACTTCATGTTAAAAACCTCAATAAATTAAGTATGCTGGACCTATCTCAAAA	160045
Qy	358	TAATAAAAAGCCATATATGACGAGCCACACAGCAAACTTATATCGAATGGGAAAAGTTGCA	417
Db	160046	TAATTAAGAGCTATCTATGTCAAACCAACGACCAATATCATCTGATGGGCAAAAACCTGG	160105
Qy	418	AAACATTTGCTCCTGAGAACTGGAACAAGACAAAGATG-CTACTTTCAACACTTCTATTC	476
Db	160106	AAGCATTTCCCTTTGAAAAATGGGACAAGACAGGAGATGCCCTCTCTCACCACTCTTATTC	160165
Qy	477	ACATAGTAGTGAAGTTTATAGCCAGGCAATCAGAAAGAAAGAAATCAAGGGACACC	536
Db	160166	ACATAGTGTGGAAGATCTGTCCAGGGCAATCAGGAGGAAGAAATTAAGGGATATTC	160225
Qy	537	AAATCAATTAAGAGAAAGTCAAACTGTCTCCTGTTCACATGATATGATTTATACCTAG	596
Db	160226	AATTGAGAAAAGAGGAAGTCAAAATGTCTCCTGTGTTGCAAGATGACATGATTTATATCTAG	160285
Qy	597	AAAAATCCCTAAAGCTCATCCAGAAAGCTCTAGAACTGATACATAAATTCAGTAAAGTTT	656
Db	160286	AAAAATCCCAATGTCTTCAGCCCAAAACCTCTTAAGCTGATTAAGCAATTCACGAAAGTCT	160345
Qy	657	CAGGATACAAACCTAAATGTACCAATCAAGTACACGTGCATACACCAACAGTGAACAAG	716
Db	160346	CAGGATACAAATTCATATGTCACAAATATCAACAGCAATTTCTTATACCCAAACAGACAAA	160405
Qy	717	CTGAGATCAAAATCAAGAACTCAAAACCTTTTCAATAGCTGTAAA-----AAAATACT	770
Db	160406	CAGAGGCAAAATTCATATGATGAATCCCAATTCACAATTTGTTCAAAAGAGATTAATTAAC	160465
Qy	771	TAAAGATATTTCTTACCAGAGAGGTGAAGACCTCTACACAGAAACATACAAACACAGC	830
Db	160466	TAGGATCCACTTTCACAGAGAGTGAAGACCTCTTTCACAGGAACTACAAACCACTGG	160525

QY	83	T----	GACATCATAGATGACACAAA	CAAGTGGAAACATCC	ATGCTCATGATGGGTA	886	
Db	160526	TCAACGAAATTA	TAAGAGATTA	CAAAACAAATGAGACATTT	CACTGCTCATGGGTAGAA	16058585	
QY	887	GAATCAATA	TTGTGAAATGACCATAT	TGCCCCAAAGCAATCTAC	AAATTCAATGCAATTC	946	
Db	160586	GAATCAATA	TTGTGAAATGAGCCAT	ACTAGCCCAAGTAATTTAC	AGATTTCAATGCAATTC	16064545	
QY	947	CCACCAAAA	ATATCATCATCTT	CACAGAACTTGAAAAA	ACAATTCCTAAATTCATA	1006	
Db	160646	CCATCAAGCTA	CCCAATTAATCTTTCTT	CACATTA-TTGGAAAAA	ATCTACTTTAAAGTTTCA	160704	
QY	1007	TGGACACACAC	CAAAAAA	AAAAAAAAA	CCCCGATATGCCAAAGCAAGCTTA	1066	
Db	160705	TGGAC-----	CAAAAAAG	CCCGATACCAAGCTCATCT	GTAGCCAAA	160750	
QY	1067	AGAACAAAT	GTGGAGCATCACTT	ACCATCTTCAATCT	TAATCTACAGGCTTAATCA	1126	
Db	160751	AGAACAAAG	GTGGAGCATCACTAC	ACTGACTTCAAGCTTA	CAAGGCTTAACGTAA	160810	
QY	1127	CCAAACAT	CATAGGSCATGATTA	AACTAGGSCATGACCAT	TGAAAAAGAGAGAG	1186	
Db	160811	CCAAACAG	TTGTGTACTGTTAC	CAAAACAGGATATGATCA	ATGAAACAGACAGC	160870	
QY	1187	ATCCAGAA	ATTAAGCCAAATTA	TTATAGCACTGATTTT	TGACAAAGCAACAAAA	1246	
Db	160871	CCTCAGAA	ATTAATGCTGCAATAT	CGAACACTGCTGATCTTT	TGACAAACCTGAGAAAA	160930	
QY	1247	TAAAGTGG	GGAAAAAGCATCT	TCATAGTAA	CAATGCTGTGATTA	TTGGCAAGCCACAT	1306
Db	160931	AGAAATGG	GGAAAAAGATTCCT	TAATTAATTAATG	TGTGGAAAACTGCTATG	CCCATAT	160990
QY	1307	GTGGAGAA	ATGAACGTGATCCCT	TGTCTCATCTTA	TACAAAAATTTGATAC	AGATGG	1366
Db	160991	GTGAGAAAG	ATGAACGTGATCCCTT	-CCTTAC	ACTTAATACAAAAATTA	TTTCAAGATG	161049
QY	1367	ATCAAGA	CTTAATCTGAGACCT	TAAACCATAAAA	TTCTAGAAATTA	CAATCAGAAA	1426
Db	161050	ATTAAAGA	CTTAATCTGAGACCT	TAAACCATAAAA	CCCTAGAAAGAAA	CTTAGGCAAT	161109
QY	1427	ATGCTCT	GTGACCTCAT	TAGGCAAGCTT	CATGGCCAAAGCC	CAAAAGTAAATGCA	1486
Db	161110	ACCATTC	AGACATGAGCA	TGGCAATGGCAAG	CTTCATGTTTAA	CAACAAAGCAATG	161169
QY	1487	ACAAAA	CAAAATAATAT	ATAGATGACCTTA	TTAACTTAA	AAAGCTTTGCGAC	1546
Db	161170	ACAAAAG	CCAAATTTGACAAAT	TGGGATCTTA	TTAACTTAAAG	ACTTCTGCA	161229
QY	1547	ACAATCA	TTAGCAGAGCA	AAACAGACCC	ACGATGAGAA	AAATCTTCA	1606
Db	161230	GAATCTA	ACCACTGAGATG	ATAGGCACT	CAAGATGGGAAAA	TTTTTGC	161289
QY	1607	GCAATCG	ATTAAGCTAATAT	TCGGAAATCC	CAAGAACTCA	AAACATCAGCA	1666
Db	161290	TCATTCG	AAAGGCTTAATTT	TCAGATCT	TCAGAAAGAACTTAA	CAAAATGTTCA	161349
QY	1667	AAAGCAA	CAATCCAT	GTAAAGATGGGCT	TAAGCATGAAT	GACATTTCTCAA	1726
Db	161350	A----	AAACAA	CCCCCATCAAAAA	ATGGGCAAAAT	CAAGAAAGCACTT	161405
QY	1727	GATATCA	AAATGGCCAA--	AAACAG	AAAAAATGCTTA	CAATGATCA	1784
Db	161406	GACATTT	ATGCAACCA	CAAGACAT	GTAAAAATGCTAT	CATCACTGGCCAT	161465
QY	1785	ATGTAAT	CAACACTG	TAATGCGAT	ACACCTTACCT	CGACAAATG	1844
Db	161466	ATGGAAT	CAAAACCA	CAATGAGAT	ATACATCTCAC	ACAGTTA	161523
QY	1845	AAATCT	TAATAATAT	TATGATGTT	GTGGGTCTGT	GTGTATAA	1904
Db	161524	AAAGTC	AGAAACAA	ACAGATGCT	GAGAGACAT	GTGAGAAATAT	161583


```

Db 577826 CATTAGACTGAACAGGCACTCAACAAGTGGGAGAAAATTTTGCATCTATTCATCTGA 577767
QY 1615 CTAAAGACTAATATTCGGATCCACAGGAATCTCAACAAATCAGCAGAGAGAACAA 1674
Db 577766 CAAAAGGCTAAATTCAGATCTACAAATGAACTCAACAAATTTACAGAGAAAATCAAA 577707
QY 1675 CAATCCCATGAAGAAGTGGGCTAAGGACATGAATTAGACAATTCTCAAAAGAAATPACA 1734
Db 577706 CAACCCCATCAAAAAGTGGGCAAAAGATGAAACAGACACTTCTCACAGAGACATTTA 577647
QY 1735 AATGGCAACAA--ACAGGAAAAAATGCTTAACATGCTAATGATTAGGAAATGTAAAT 1792
Db 577646 TGCAGCCAAAAACACAGAAAAAATGCTCATCATCTGGCCATCAGAGAAATGCAAT 577587
QY 1793 CAACCTGTAAATGCGATACCACTTACTCTGCAAGATGTCATTAATTTAAATCTAA 1852
Db 577586 CAAAACCAATGGGATACCATCTCAACACAGTTAGATGGCGATCATTAATAAAGTCAGG 577527
QY 1853 AAATAATAGATGTGGTGGGTCTGTGTGATAAAGAACCTTTACACTGCTGTGGGA 1912
Db 577526 AAACHACAGGTCTGGAGAGATGTGAGAAATAGAAATCTTTTCACTGTGTGGGA 577467
QY 1913 ATGTAACCTTGGCAACCACTATGAAAAAGTGTGAAATTTCTTAAGGAACTAAAGT 1972
Db 577466 CTGTAACTAGTTCACACATGTGGAAGTCAAGTGTGATTTCTTCAGGATCTAGAACT 577407
QY 1973 AGATGACCATTTGATCCAGCAATCCCAATTAATATGATATAAATATATATTTATATAC 2032
Db 577406 AGAATATACCATTTGACCCAGATCCCATTT-ACTGGGTATATACCCAAAGATTAATAAT 577348
QY 2033 CATG 2036
Db 577347 CATG 577344

```

Search completed: December 2, 2004, 18:17:25
 Job time : 730 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 2, 2004, 01:15:24 ; Search time 4225 Seconds
(without alignments)
1123.321 Million cell updates/sec

Title: US-09-867-570-3

Perfect score: 8622
Sequence: 1 tgratgaagcgaatgctcact.....gtccccaagcccttacc 8622

Scoring table:
IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3694831 seqs, 2747406616 residues

Total number of hits satisfying chosen parameters: 7389662

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8622	100.0	8622	11	US-09-867-570-3
2	1429.4	16.6	398287	17	US-10-741-601-5719
3	1401	16.2	66494	17	US-10-450-826-47
4	1395	16.2	209484	13	US-10-087-192-418
5	1395	16.2	209484	18	US-10-331-053-4
6	1390.4	16.1	465237	9	US-09-933-267A-1
7	1375.6	16.0	1691139	14	US-10-067-514-1
8	1375.6	15.9	1691139	16	US-10-419-723-1
9	1374.6	15.9	1400	14	US-10-183-116-30
10	1374.6	15.9	1400	15	US-10-225-567A-673
11	1370	15.9	100944	17	US-10-322-696-4
12	1366.4	15.8	277616	17	US-10-367-094-83

C	13	1358	15.8	99014	9	US-09-880-107-3428	Sequence 3428, Ap
	14	1346	15.6	136284	17	US-10-775-169-149	Sequence 149, Ap
	15	1311.6	15.2	1503841	9	US-09-795-668-1	Sequence 1, Appl1
	16	1311.6	15.2	1503841	9	US-09-795-668-1	Sequence 1, Appl1
	17	1311.6	15.2	1503841	9	US-09-946-807-1	Sequence 1, Appl1
	18	1308	15.2	358246	15	US-10-292-798-1095	Sequence 1095, Ap
	19	1302.6	15.1	81098	13	US-10-087-192-2032	Sequence 2032, Ap
	20	1302	15.1	145608	17	US-10-721-693-7	Sequence 7, Appl1
	21	1302	15.1	145608	18	US-10-852-997-7	Sequence 7, Appl1
	22	1301.8	15.1	32167	10	US-09-764-891-8197	Sequence 8197, Ap
	23	1289	15.0	290547	17	US-10-367-094-77	Sequence 77, Appl1
	24	1288.6	14.9	247682	16	US-10-235-192A-28	Sequence 28, Appl1
	25	1286.8	14.9	148567	9	US-09-801-8768-3	Sequence 3, Appl1
	26	1286.8	14.9	148567	14	US-10-254-863-3	Sequence 3, Appl1
	27	1286.8	14.9	148567	17	US-10-667-442-3	Sequence 166, Appl
	28	1286.8	14.9	59475	16	US-10-322-696-166	Sequence 1273, Ap
	29	1274	14.8	1369	15	US-10-292-798-1273	Sequence 8024, Ap
	30	1271.6	14.7	32249	10	US-09-764-891-8024	Sequence 32, Appl
	31	1264	14.7	1604	14	US-10-183-116-32	Sequence 68, Appl
	32	1258.2	14.6	1604	15	US-10-225-567A-688	Sequence 58, Appl
	33	1258.2	14.6	246940	17	US-10-322-696-58	Sequence 2002, Ap
	34	1247.8	14.5	261817	13	US-10-087-192-2002	Sequence 5611, Ap
	35	1247.4	14.5	253861	17	US-10-741-601-5611	Sequence 40, Appl
	36	1247	14.5	79684	15	US-10-034-650-40	Sequence 149, Ap
	37	1245.6	14.4	136284	17	US-10-775-169-149	Sequence 5719, Ap
	38	1234.8	14.3	398287	17	US-10-741-601-5719	Sequence 15, Appl
	39	1228.4	14.2	2040	14	US-10-183-116-15	Sequence 427, Ap
	40	1225.8	14.2	267156	11	US-09-968-007A-427	Sequence 154, Ap
	41	1225.8	14.2	684973	9	US-09-263-959-1	Sequence 28, Appl
	42	1219.8	14.1	49536	13	US-10-087-192-154	Sequence 1059, Ap
	43	1219.8	14.1	49536	18	US-10-331-053-28	Sequence 10, Appl
	44	1208.8	14.0	21188	15	US-10-292-798-1059	
	45	1203	14.0	130505	17	US-10-367-094-10	

ALIGNMENTS

RESULT 1
US-09-867-570-3
Sequence 3, Application US/09867570
Publication No. US20040076951A1
GENERAL INFORMATION:
APPLICANT: WEI, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
FILE REFERENCE: C1000900-CIP
CURRENT APPLICATION NUMBER: US/09/867,570
CURRENT FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 09/695,045
PRIOR FILING DATE: 2000-10-25
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 8622
TYPE: DNA
ORGANISM: Human
US-09-867-570-3

Query Match 100.0%; Score 8622; DB 11; Length 8622;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 8622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	TGATGAGCAGCATGCTTATATACCAACGAGAAAGATATCAAAAAAGAAAC	60
DB	1	TGATGAGCAGCATGCTTATATACCAACGAGAAAGATATCAAAAAAGAAAC	60
QY	61	TATAGACCAATACCATGATGATATATACGAGAAATCCCAAAATATAGCTTAC	120
DB	61	TATAGACCAATACCATGATGATATATACGAGAAATCCCAAAATATAGCTTAC	120
QY	121	CCATCCACAGCATATCAAGAAATATCCACCATTTGTCAAGTGGTTTCATACACGG	180

Db 121 CCAATCCAAAGCATATCAAGAAAGTAATCCACCAATTCCTCAATGGGTTTCATACCAAGG 180
 Qy 181 GTGCGAGATAGGTTAAATATCACAAGTCAATAATATGTATATCATCAATTAACAGAAATTA 240
 Db 181 GTGCGAGATAGGTTAAATATCACAAGTCAATAATATGTATATCATCAATTAACAGAAATTA 240
 Qy 241 AAAAATAAATATCATATGATCATCTGATATGATGCTGTAAGAAAGAAATTTGACAAATTTAA 300
 Db 241 AAAAATAAATATCATATGATCATCTGATATGATGCTGTAAGAAAGAAATTTGACAAATTTAA 300
 Qy 301 ATTTCTTTATGATTAATAAATCTTCAGCAAAATCGACATAGAAAGACATATTAATGTA 360
 Db 301 ATTTCTTTATGATTAATAAATCTTCAGCAAAATCGACATAGAAAGACATATTAATGTA 360
 Qy 361 TAAAAGCCATATATGACGGAACCCACAGCAAACTTATGTAATGAGGAAAGTTGAAAA 420
 Db 361 TAAAAGCCATATATGACGGAACCCACAGCAAACTTATGTAATGAGGAAAGTTGAAAA 420
 Qy 421 CATTTCCCTGAGAACTGGAACAAGACAGAGATGCTACTTCACTTCTATTCACAT 480
 Db 421 CATTTCCCTGAGAACTGGAACAAGACAGAGATGCTACTTCACTTCTATTCACAT 480
 Qy 481 AGTAGTGAAATTTAGCCAGAGCATCAGACAGAGAAAGAAATGAGGCAACCCAAAT 540
 Db 481 AGTAGTGAAATTTAGCCAGAGCATCAGACAGAGAAAGAAATGAGGCAACCCAAAT 540
 Qy 541 CAATTAAGAGAAATGCAAACTGTCCCTGTTCACTGATGATATGATGTAATCTTAAAGAA 600
 Db 541 CAATTAAGAGAAATGCAAACTGTCCCTGTTCACTGATGATATGATGTAATCTTAAAGAA 600
 Qy 601 CCTTAAGACTCATCCAGAAAGCTCTAGAACTGATACATTAATTCAGTAAAGTTTCAGG 660
 Db 601 CCTTAAGACTCATCCAGAAAGCTCTAGAACTGATACATTAATTCAGTAAAGTTTCAGG 660
 Qy 661 ATACAAACTAAATGTACACAAATCAGTAGACCTGCTATACACCAAGAGTGAACCAAGCTGA 720
 Db 661 ATACAAACTAAATGTACACAAATCAGTAGACCTGCTATACACCAAGAGTGAACCAAGCTGA 720
 Qy 721 GAATCAATCAAGAACTCAAACTTTTACATATGCTGTAAGAAATTAATCTTAAGAAATTT 780
 Db 721 GAATCAATCAAGAACTCAAACTTTTACATATGCTGTAAGAAATTAATCTTAAGAAATTT 780
 Qy 781 CTATCCAAAGAGGTGAAGAGCTCTACAAAGAAATCTACAAACAGGTGACATCATTA 840
 Db 781 CTATCCAAAGAGGTGAAGAGCTCTACAAAGAAATCTACAAACAGGTGACATCATTA 840
 Qy 841 GATGACACAAAGAGTGAAGACATCCCATGCTCATGATGAGGTGAATCAATATTTG 900
 Db 841 GATGACACAAAGAGTGAAGACATCCCATGCTCATGATGAGGTGAATCAATATTTG 900
 Qy 901 AAAATGACATATTTGCAAAAGCAATCTACAGTTTCAATGCAATTTCCACCAAAATATCA 960
 Db 901 AAAATGACATATTTGCAAAAGCAATCTACAGTTTCAATGCAATTTCCACCAAAATATCA 960
 Qy 961 TCATCATTTCTTACAGAACTAGAAAAAAACAATTTCTAAATTCATATGGAACAACA 1020
 Db 961 TCATCATTTCTTACAGAACTAGAAAAAAACAATTTCTAAATTCATATGGAACAACA 1020
 Qy 1021 AAAAAAATAAATAAATCCGCAATAGCCAAAGCAAGCTTAGCAAAAAAGAACAAATCTGGA 1080
 Db 1021 AAAAAAATAAATAAATCCGCAATAGCCAAAGCAAGCTTAGCAAAAAAGAACAAATCTGGA 1080
 Qy 1081 GGCATCATCTTACCTTTCAAACTATATCTACAGGCTTAAATCACCACAAACATCATG 1140
 Db 1081 GGCATCATCTTACCTTTCAAACTATATCTACAGGCTTAAATCACCACAAACATCATG 1140
 Qy 1141 CACTGACATATAAAGTGAAGCATAGACCAATGGAAGAAAGAGAAATCCAGAAATTAAG 1200
 Db 1141 CACTGACATATAAAGTGAAGCATAGACCAATGGAAGAAAGAGAAATCCAGAAATTAAG 1200
 Qy 1201 CCAAAATATTAATAGCCAACTGATTTTGAACAAAGCAAAACAAATTAAGGTGGGAAAA 1260

Db 1201 CCAAAATATTAATAGCAACTGATTTTGAACAAAGCAAAACAAATTAAGTGGGAAAA 1260
 Qy 1261 GACATTTCTATGTTAACAATATGCTGAGATTTTGGCAAGCCACATGTGGAAGAAATGAAA 1320
 Db 1261 GACATTTCTATGTTAACAATATGCTGAGATTTTGGCAAGCCACATGTGGAAGAAATGAAA 1320
 Qy 1321 CTGGATCCCTGTGCTCTCACTTAATACAAATATGATACAGATGATCAAAAGCTTAAA 1380
 Db 1321 CTGGATCCCTGTGCTCTCACTTAATACAAATATGATACAGATGATCAAAAGCTTAAA 1380
 Qy 1381 TCTGAGACCTTAACCCATTAATAATTTCTAGAGATTAATCAGAAAAATGCTTACAT 1440
 Db 1381 TCTGAGACCTTAACCCATTAATAATTTCTAGAGATTAATCAGAAAAATGCTTACAT 1440
 Qy 1441 TCACTTAGCAAGAACTTATGAGCCAGAAACCCAAAGTAAATGGAACAAAAACAAAT 1500
 Db 1441 TCACTTAGCAAGAACTTATGAGCCAGAAACCCAAAGTAAATGGAACAAAAACAAAT 1500
 Qy 1501 AATATAGTGAAGCTTAATTAATTAATAAAGCTTTTGGCAGCAAAACATCATTAAGCAG 1560
 Db 1501 AATATAGTGAAGCTTAATTAATTAATAAAGCTTTTGGCAGCAAAACATCATTAAGCAG 1560
 Qy 1561 AGCAAAACAGAACCCACCGAGTGAAGAAATCTTCAAACTAAGCATCTGACTAAGG 1620
 Db 1561 AGCAAAACAGAACCCACCGAGTGAAGAAATCTTCAAACTAAGCATCTGACTAAGG 1620
 Qy 1621 ACTTAATATCCGGAATTCACAAAGAACTCAAAATTCAGCAAGAAAGAAACAAATCC 1680
 Db 1621 ACTTAATATCCGGAATTCACAAAGAACTCAAAATTCAGCAAGAAAGAAACAAATCC 1680
 Qy 1681 CATGAAGAGTGGGCTTAAGACATGAATAGCAATTCACAAAGAAAGATATCAATGGC 1740
 Db 1681 CATGAAGAGTGGGCTTAAGACATGAATAGCAATTCACAAAGAAAGATATCAATGGC 1740
 Qy 1741 CAACAAACAGAAAAAATGCTTAACATCACTAATATTAAGGAAATGTAAATCAACCTG 1800
 Db 1741 CAACAAACAGAAAAAATGCTTAACATCACTAATATTAAGGAAATGTAAATCAACCTG 1800
 Qy 1801 TAAATGGAATACCACTTACTCTGCAAGAAATGCTAATTTAAATCTAATAATATA 1860
 Db 1801 TAAATGGAATACCACTTACTCTGCAAGAAATGCTAATTTAAATCTAATAATATA 1860
 Qy 1861 GATGTTGGTGGGCTGTGTGTGATTAAGAAACCTTTACACGCTGTGGGAAATTAAC 1920
 Db 1861 GATGTTGGTGGGCTGTGTGTGATTAAGAAACCTTTACACGCTGTGGGAAATTAAC 1920
 Qy 1921 TTGGGCAACCACTATGAAAAACAGTGTGAAATTTCTTAAGAACTAAGATGATCAAC 1980
 Db 1921 TTGGGCAACCACTATGAAAAACAGTGTGAAATTTCTTAAGAACTAAGATGATCAAC 1980
 Qy 1981 CATTTGATCAGCAATCCCATTAATATGTAATTAATTAATTAATTAATTAATTAAT 2040
 Db 1981 CATTTGATCAGCAATCCCATTAATATGTAATTAATTAATTAATTAATTAATTAAT 2040
 Qy 2041 ACAACTCAGCCATTAATAAATAAATGATGACATTCACAGCAATCTAGATGAATGG 2100
 Db 2041 ACAACTCAGCCATTAATAAATAAATGATGACATTCACAGCAATCTAGATGAATGG 2100
 Qy 2101 AGACCTTAATTTCTAAGTGGGTAATCTCAGAAATGAAAAACAAACATCATATGTTCTAC 2160
 Db 2101 AGACCTTAATTTCTAAGTGGGTAATCTCAGAAATGAAAAACAAACATCATATGTTCTAC 2160
 Qy 2161 TTAAGATGAGGCTTAAGCTGTGAGAGACGAAAGCATTAAGTGAATTAATGAATCTGG 2220
 Db 2161 TTAAGATGAGGCTTAAGCTGTGAGAGACGAAAGCATTAAGTGAATTAATGAATCTGG 2220
 Qy 2221 GGAATGAGGGAAGAGATGGAAGAGAGGCGAAGGTAATAAGACTACACATGAGTACAGT 2280
 Db 2221 GGAATGAGGGAAGAGATGGAAGAGAGGCGAAGGTAATAAGACTACACATGAGTACAGT 2280
 Qy 2281 GTAACCTGCTCAGGTGATGAGGTGACCAAAATCTCAGAAATTAACCACTTAAGAACTTATC 2340
 Db 2281 GTAACCTGCTCAGGTGATGAGGTGACCAAAATCTCAGAAATTAACCACTTAAGAACTTATC 2340

QY 2341 CATGGAGCAAAACACACTCTTCCCAAAATCCCAATGAATATAATTAATTAATA 2400
DB 2341 CATGGAGCAAAACACACTCTTCCCAAAATCCCAATGAATATAATTAATTAATA 2400
QY 2401 ATGATTTAATTTTCAGAGATTTTAAAAAGTTCAGCTGTTCAGAGTTTAAATTAAGAGTA 2460
DB 2401 ATGATTTAATTTTCAGAGATTTTAAAAAGTTCAGCTGTTCAGAGTTTAAATTAAGAGTA 2460
QY 2461 AGAATGAAAAGTGTAGCAAGTGTAGCTGTGACAAATGGGACTCTAGATTTTACCTTG 2520
DB 2461 AGAATGAAAAGTGTAGCAAGTGTAGCTGTGACAAATGGGACTCTAGATTTTACCTTG 2520
QY 2521 CATACACTTCTCTGGCAATTTTGAAGAAAGTATACATGAATATATACCACTATATGA 2580
DB 2521 CATACACTTCTCTGGCAATTTTGAAGAAAGTATACATGAATATATACCACTATATGA 2580
QY 2581 AAGAAAATCATCAAAAATGTGTGACAGGCAATGTGACGCTTGAATGGTCCCATGATCTAC 2640
DB 2581 AAGAAAATCATCAAAAATGTGTGACAGGCAATGTGACGCTTGAATGGTCCCATGATCTAC 2640
QY 2641 TTTTTCATTTGGATATTAAGCCTCATATGATGATTCATATGCTTAATGTAGCTTAG 2700
DB 2641 TTTTTCATTTGGATATTAAGCCTCATATGATGATTCATATGCTTAATGTAGCTTAG 2700
QY 2701 GCCCATATTTGATTTTAAATTCAGACAGCAATTAATTACAGAGAGTTGAACAAGATGG 2760
DB 2701 GCCCATATTTGATTTTAAATTCAGACAGCAATTAATTACAGAGAGTTGAACAAGATGG 2760
QY 2761 GACGTATAGAGAGGCTTAAATGTATCTGATATGAGGACAGAGGCAAGATCATCTCAG 2820
DB 2761 GACGTATAGAGAGGCTTAAATGTATCTGATATGAGGACAGAGGCAAGATCATCTCAG 2820
QY 2821 TTAGATTTTGTCTCAATTAATCTCTGACCTCTGATTTGCCCATAGTCTTATACAGGAA 2880
DB 2821 TTAGATTTTGTGTCTCAAAATACCTCTGACCTCTGATTTGCCCATAGTCTTATACAGGAA 2880
QY 2881 ATTAACAAGACTGTACAGATTTTGTATAGCCTGATGCTGACCAAGCTTTTATTCAGCT 2940
DB 2881 ATTAACAAGACTGTACAGATTTTGTATAGCCTGATGCTGACCAAGCTTTTATTCAGCT 2940
QY 2941 CCTGTAGGACATCTCTGAATTAAGCAACAGAAAAGCTCTGAAGTCACTGAATCCCA 3000
DB 2941 CCTGTAGGACATCTCTGAATTAAGCAACAGAAAAGCTCTGAAGTCACTGAATCCCA 3000
QY 3001 GAAAGGCTCTCTACCTTTAGCAAGGAGGCTTCTCACTGAGCAAAAGAGAGAGAT 3060
DB 3001 GAAAGGCTCTCTACCTTTAGCAAGGAGGCTTCTCACTGAGCAAAAGAGAGAGAT 3060
QY 3061 AAGGATAGTACCAAGAACTCTCTTCTTCCACAGTCAAGTATGATTTTGTGTAGATC 3120
DB 3061 AAGGATAGTACCAAGAACTCTCTTCTTCCACAGTCAAGTATGATTTTGTGTAGATC 3120
QY 3121 ATGTCTTATGCTTCCACTTGTGTCTACATGACGAGGAGTCAAGAGCTGTGTCAGAAA 3180
DB 3121 ATGTCTTATGCTTCCACTTGTGTCTACATGACGAGGAGTCAAGAGCTGTGTCAGAAA 3180
QY 3181 AGACAGAGACATGAGCTTCTTTCAGAAAATGAGTGTCTGCAACCAAACTGTGTAG 3240
DB 3181 AGACAGAGACATGAGCTTCTTTCAGAAAATGAGTGTCTGCAACCAAACTGTGTAG 3240
QY 3241 CTCTAATATGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3300
DB 3241 CTCTAATATGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3300
QY 3301 TTCAATCTGCTCATGAGAGACTGCTGCTCTTCAATTCATTTTAAAGAGCAAGGAGAC 3360
DB 3301 TTCAATCTGCTCATGAGAGACTGCTGCTCTTCAATTCATTTTAAAGAGCAAGGAGAC 3360
QY 3361 ATGACAGCAATTTCTTCCAGAGTTGAACGTCTATAGAGCCAGTTTCTTGTTCATCTTAC 3420
DB 3361 ATGACAGCAATTTCTTCCAGAGTTGAACGTCTATAGAGCCAGTTTCTTGTTCATCTTAC 3420

QY 3421 TTTTCAATTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3480
DB 3421 TTTTCAATTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3480
QY 3481 ATATATGAGAGTGTTCATGAGAGAAAGTATTTTCAAGATATGACAGAGTTTAAAGGTCAG 3540
DB 3481 ATATATGAGAGTGTTCATGAGAGAAAGTATTTTCAAGATATGACAGAGTTTAAAGGTCAG 3540
QY 3541 GATCTTAAGCAACAATACATAGGAGACACTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 3600
DB 3541 GATCTTAAGCAACAATACATAGGAGACACTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 3600
QY 3601 GAAAGACAGTATACATTTGTGCGCATAGCCATAGCTTACGCTTAAAGGTGGAGAGCAT 3660
DB 3601 GAAAGACAGTATACATTTGTGCGCATAGCCATAGCTTACGCTTAAAGGTGGAGAGCAT 3660
QY 3661 GAGAGGCAAGTGAAGAAAGCCTGCGAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGT 3720
DB 3661 GAGAGGCAAGTGAAGAAAGCCTGCGAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGT 3720
QY 3721 TGATCTGTGTCTTCCACCAAAATCTCATGTGATTTGTAATTTCCAAATGTTGAGAGAGG 3780
DB 3721 TGATCTGTGTCTTCCACCAAAATCTCATGTGATTTGTAATTTCCAAATGTTGAGAGAGG 3780
QY 3781 CTTGTGGAGATGATATTTATGATCAAGGAGATGTTTGTGATGATATTTTAAACCAT 3840
DB 3781 CTTGTGGAGATGATATTTATGATCAAGGAGATGTTTGTGATGATATTTTAAACCAT 3840
QY 3841 CCCCCTTGGATGTTGTTGTGATCTGACAGAGTCTCATGAAATCTAGTTGTTTAAA 3900
DB 3841 CCCCCTTGGATGTTGTTGTGATCTGACAGAGTCTCATGAAATCTAGTTGTTTAAA 3900
QY 3901 GCGTGTAGCACTTCT 3960
DB 3901 GCGTGTAGCACTTCT 3960
QY 3961 CCCCCTTGGCTTCAACAGATTTGGAAGCTTCTGAGGCTCTCCCAAGAGCAAGAGCTGC 4020
DB 3961 CCCCCTTGGCTTCAACAGATTTGGAAGCTTCTGAGGCTCTCCCAAGAGCAAGAGCTGC 4020
QY 4021 TATGCTCTTGTACAGCTGTGAGAGCTATTAGCAAGTAAACCAATCTCTCTCTCTCTCT 4080
DB 4021 TATGCTCTTGTACAGCTGTGAGAGCTATTAGCAAGTAAACCAATCTCTCTCTCTCTCT 4080
QY 4081 TCCAGTCTCAGGATTTCTTTTTCAGAAATTTGAGATGAACTAATACAGACAGAGAG 4140
DB 4081 TCCAGTCTCAGGATTTCTTTTTCAGAAATTTGAGATGAACTAATACAGACAGAGAG 4140
QY 4141 CAGAGATGAGAAATCCCAAGGAGTCTTCTGCTGTCTTCAAGTCTCTCTGCTGCTC 4200
DB 4141 CAGAGATGAGAAATCCCAAGGAGTCTTCTGCTGTCTTCAAGTCTCTCTGCTGCTC 4200
QY 4201 CCAAGTCTCAATTTCCACAGAAACAGAAATTAAGAAATCCCAAGTGTGTATATA 4260
DB 4201 CCAAGTCTCAATTTCCACAGAAACAGAAATTAAGAAATCCCAAGTGTGTATATA 4260
QY 4261 GAAAGCACTCTTGGAGTCAAAACAGATTAAGAAATGAAAGCAAAATCTCTCATGG 4320
DB 4261 GAAAGCACTCTTGGAGTCAAAACAGATTAAGAAATGAAAGCAAAATCTCTCATGG 4320
QY 4321 TAAATGAGATATCCCTCTCACTCTTGTATCTCTTAATTTCTGAGGCTTCTCTATC 4380
DB 4321 TAAATGAGATATCCCTCTCACTCTTGTATCTCTTAATTTCTGAGGCTTCTCTATC 4380
QY 4381 TGAATATCCCTGTCTCATTTTCACTTATCAGACTATCAGACTTATTTTGTGCTTCTTC 4440
DB 4381 TGAATATCCCTGTCTCATTTTCACTTATCAGACTATCAGACTTATTTTGTGCTTCTTC 4440
QY 4441 TCTACTGTCACTTTTATGAGAAATGTTTGCATTTTGTAAAAATGATAGAAATTAAT 4500
DB 4441 TCTACTGTCACTTTTATGAGAAATGTTTGCATTTTGTAAAAATGATAGAAATTAAT 4500
QY 4501 GTAAATTTAAAAAGACATATGATTTTGTGTAAGAAATTAAGTTTGGCTATCTAATAA 4560

Db 4501 GTAAATTTAAAGAAACATATGATTTTGTGTAATTAAGTTTGCTGATTAATTA 4560
Qy 4561 GACATGAAGAAATATCTTAAACAGAAAGTATAGTTGCTTGCTGCTGCTAGGTT 4620
Db 4561 GACATGAAGAAATATCTTAAACAGAAAGTATAGTTGCTTGCTGCTGCTAGGTT 4620
Qy 4621 CTGAATCTACAGATTCACAACTACAGAGGAACTTTTCCAAAATAAAGGTGGCG 4680
Db 4621 CTGAATCTACAGATTCACAACTACAGAGGAACTTTTCCAAAATAAAGGTGGCG 4680
Qy 4681 GAGTTGTATCTAGAACAGGTACAACTTTATTTCTTTGTCATTAATTTCTGAAAA 4740
Db 4681 GAGTTGTATCTAGAACAGGTACAACTTTATTTCTTTGTCATTAATTTCTGAAAA 4740
Qy 4741 CTACATATTAACAGAACTTATATAGCAATTTGCAATTTGTCAGTTATTTCTAA 4800
Db 4741 CTACATATTAACAGAACTTATATAGCAATTTGCAATTTGTCAGTTATTTCTAA 4800
Qy 4801 AAATGATTAATGATCTGGGAGAAAGTGCATGATTAATTAACATTAATTAAG 4860
Db 4801 AAATGATTAATGATCTGGGAGAAAGTGCATGATTAATTAACATTAATTAAG 4860
Qy 4861 AAATGAGCATCTGCAAGATTTTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 4920
Db 4861 AAATGAGCATCTGCAAGATTTTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 4920
Qy 4921 ACACAAAATGACACTCTTGAGATCTGAACTGAAGCTCCAAAGCATCATACAGAA 4980
Db 4921 ACACAAAATGACACTCTTGAGATCTGAACTGAAGCTCCAAAGCATCATACAGAA 4980
Qy 4981 TTCCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5040
Db 4981 TTCCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5040
Qy 5041 GGTTCACAGCAGATTGAGATTCAGATCTTAATGAAAAAGAGAGAGAGAGGCT 5100
Db 5041 GGTTCACAGCAGATTGAGATTCAGATCTTAATGAAAAAGAGAGAGAGAGGCT 5100
Qy 5101 TTGCTCTCTTAATTAATCCCAAGAGCAGAGACTGCTGCTGCTGCTGCTGCTGCT 5160
Db 5101 TTGCTCTCTTAATTAATCCCAAGAGCAGAGACTGCTGCTGCTGCTGCTGCTGCT 5160
Qy 5161 ACTTAACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5220
Db 5161 ACTTAACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5220
Qy 5221 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5280
Db 5221 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5280
Qy 5281 ACCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5340
Db 5281 ACCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5340
Qy 5341 GGAACCTCTCAGAAATTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5400
Db 5341 GGAACCTCTCAGAAATTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5400
Qy 5401 CTCGAGGAGCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5460
Db 5401 CTCGAGGAGCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5460
Qy 5461 TTAATTTCTAGAGCTTACCTTATGTTATGAGGATTAATTAACATTAAGCTCT 5520
Db 5461 TTAATTTCTAGAGCTTACCTTATGTTATGAGGATTAATTAACATTAAGCTCT 5520
Qy 5521 TTAAGAAATCATGATGATTAATTAACATTAATGCTGCTGCTGCTGCTGCTGCT 5580
Db 5521 TTAAGAAATCATGATGATTAATTAACATTAATGCTGCTGCTGCTGCTGCTGCT 5580
Qy 5581 AGACAGCATGATTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5640
Db 5581 AGACAGCATGATTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5640

Db 5581 AGACAGCATGATTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5640
Qy 5641 TACCATCTTAATGATGACAAACGTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 5700
Db 5641 TACCATCTTAATGATGACAAACGTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 5700
Qy 5701 CTTTCTCTCCATTTCTTTTGGGCAATCAGAGCTGCTGCTGCTGCTGCTGCTGCT 5760
Db 5701 CTTTCTCTCCATTTCTTTTGGGCAATCAGAGCTGCTGCTGCTGCTGCTGCTGCT 5760
Qy 5761 AGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5820
Db 5761 AGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5820
Qy 5821 TTCTTTCCACCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5880
Db 5821 TTCTTTCCACCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5880
Qy 5881 CCTGATGAGCAAAAGAACTGCTGATTTGTTCTTAATGAGCAGCAGCTGCTGCTGCT 5940
Db 5881 CCTGATGAGCAAAAGAACTGCTGATTTGTTCTTAATGAGCAGCAGCTGCTGCTGCT 5940
Qy 5941 TTTTCCCTTAATTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6000
Db 5941 TTTTCCCTTAATTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6000
Qy 6001 TTTTTCAGCCCAATTTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6060
Db 6001 TTTTTCAGCCCAATTTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6060
Qy 6061 TTAACCCCTCTTAACATGATTAATGATGATGATGATGATGATGATGATGATGATGAT 6120
Db 6061 TTAACCCCTCTTAACATGATTAATGATGATGATGATGATGATGATGATGATGATGAT 6120
Qy 6121 AAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6180
Db 6121 AAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6180
Qy 6181 CCGGCAAAATTAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6240
Db 6181 CCGGCAAAATTAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6240
Qy 6241 TACTAATTTTAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6300
Db 6241 TACTAATTTTAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6300
Qy 6301 TCAATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6360
Db 6301 TCAATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6360
Qy 6361 GCATGAATATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6420
Db 6361 GCATGAATATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6420
Qy 6421 CCTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6480
Db 6421 CCTCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6480
Qy 6481 AATGAATAGTTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6540
Db 6481 AATGAATAGTTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6540
Qy 6541 TCCCAACCTTCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6600
Db 6541 TCCCAACCTTCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6600
Qy 6601 CTGTTCTAATGCTGAGCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6660
Db 6601 CTGTTCTAATGCTGAGCAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6660
Qy 6661 ATGCTTTCTAGAGGCTTGGGCAAGCACTGATGCTTATTTTAAAGGCACTAGTGT 6720
Db 6661 ATGCTTTCTAGAGGCTTGGGCAAGCACTGATGCTTATTTTAAAGGCACTAGTGT 6720

QY 6721 CAATTTGAGTTTGGGCACTCATCATCATCTTCTCAACAAGATAGAGCTGCCA 6780
 Db 6721 CAATTTGAGTTTGGGCACTCATCATCATCTTCTCAACAAGATAGAGCTGCCA 6780
 QY 6781 AATGAAATTCGATGAAATGAAATTTCTCATCTAATATATATGATGATCTAATGCT 6840
 Db 6781 AATGAAATTCGATGAAATGAAATTTCTCATCTAATATATATGATGATCTAATGCT 6840
 QY 6841 TACATGTCCTTCAATTTTATTTTCAATTTCAATCAATCTACATTCATTTAGGCTT 6900
 Db 6841 TACATGTCCTTCAATTTTATTTTCAATTTCAATCAATCTACATTTAGGCTT 6900
 QY 6901 CTGATGATGATCTTCTCATTTGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 6960
 Db 6901 CTGATGATGATCTTCTCATTTGAAATGAAATGAAATGAAATGAAATGAAATGAAATG 6960
 QY 6961 GAATGAGCACTGAGGATTAATAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 7020
 Db 6961 GAATGAGCACTGAGGATTAATAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 7020
 QY 7021 GAGTTATTTCTCAACCAAGAGTCCGCAATTTTCAACATCTTACGGAAGATCCATTT 7080
 Db 7021 GAGTTATTTCTCAACCAAGAGTCCGCAATTTTCAACATCTTACGGAAGATCCATTT 7080
 QY 7081 TCCCTGCTGATTAATGATGATCTTTTCTCTGCAAGATGACCAATCTGATGATG 7140
 Db 7081 TCCCTGCTGATTAATGATGATCTTTTCTCTGCAAGATGACCAATCTGATGATG 7140
 QY 7141 AGGGTGCACAAACCACTTTGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 7200
 Db 7141 AGGGTGCACAAACCACTTTGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 7200
 QY 7201 CCCAGATAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7260
 Db 7201 CCCAGATAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7260
 QY 7261 CTTAATTTATCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7320
 Db 7261 CTTAATTTATCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7320
 QY 7321 AGAGTCAACAAAGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 7380
 Db 7321 AGAGTCAACAAAGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 7380
 QY 7381 CTTAATTTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7440
 Db 7381 CTTAATTTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7440
 QY 7441 CAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7500
 Db 7441 CAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7500
 QY 7501 AGGGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7560
 Db 7501 AGGGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7560
 QY 7561 CTGACACCAATCAACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7620
 Db 7561 CTGACACCAATCAACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7620
 QY 7621 GGGCTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7680
 Db 7621 GGGCTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7680
 QY 7681 CTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7740
 Db 7681 CTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7740
 QY 7741 GACTTCTCTTCTTACGAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATG 7800
 Db 7741 GACTTCTCTTCTTACGAGCAATTAATGATGATGATGATGATGATGATGATGATGATGATG 7800

QY 7801 CATCCCATCTCCAAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 7860
 Db 7801 CATCCCATCTCCAAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 7860
 QY 7861 ATGCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7920
 Db 7861 ATGCTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7920
 QY 7921 TGCCGCGCGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7980
 Db 7921 TGCCGCGCGCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7980
 QY 7981 CTGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8040
 Db 7981 CTGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8040
 QY 8041 TGCTGTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8100
 Db 8041 TGCTGTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8100
 QY 8101 TGCTGTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8160
 Db 8101 TGCTGTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8160
 QY 8161 ACCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8220
 Db 8161 ACCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8220
 QY 8221 TTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8280
 Db 8221 TTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8280
 QY 8281 CATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8340
 Db 8281 CATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8340
 QY 8341 TACTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8400
 Db 8341 TACTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8400
 QY 8401 CAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8460
 Db 8401 CAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8460
 QY 8461 ACCCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8520
 Db 8461 ACCCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8520
 QY 8521 AGAGCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8580
 Db 8521 AGAGCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8580
 QY 8581 TTTGCTCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8640
 Db 8581 TTTGCTCTGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8640

RESULT 2

; US-10-741-601-5719
 ; Sequence 5719, Application US/10741601
 ; Publication No. US2004016519A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CARGILL, Michele et al.
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
 ; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001500
 ; CURRENT APPLICATION NUMBER: US/10/741,601
 ; NUMBER OF SEQ ID NOS: 26415
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 5719
 ; LENGTH: 398287
 ; TYPE: DNA

ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (1)...(398287)
 OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
 US-10-741-601-5719)

Query Match 16.6%; Score 1429.4; DB 17; Length 398287;
 Identical Similarity 79.5%; Pred. No. 0;
 Matches 2004; Conservative 0; Mismatches 356; Indels 161; Gaps 20;

1 TGTATGAAGCCATGCTACTTAAATACCAAAACCAAGAAAGATATA--CAAAAAGAAA 58
 Db 331967 TCTCTGAAAGTGTGTGTCACCCCTTAATCCAAACCGGAGGACATACAAAAAGAAA 332026
 Oy 59 ACTATAGACAGTACCATGATGAATATATACAGAGAAATCCCAACAAATATCTAGCTA 118
 Db 332027 ACTACAGACCAACATCCCTGATGAACATGGGTGCAAAAATTTCAACAATTTCTAGTGA 332086
 Oy 119 ACCCAATCCAAAGCATATCAGAGAAATATCCACATTTGTCAAGGGTTTCATACCGAG 178
 Db 332087 ACTGATCCAAAGATATATCAAAAAGATTAACCTACCATGATCAAGTGATTTCTATACCG 332146
 Oy 179 GGGTCCAGG-ATAGGTTAATATACACAGTCAATTAATGTATATACATCAATTAACAGAA 237
 Db 332147 GAGTACAGAAATGATTTAATCATGTGCAAGTCAATTAATGTATATCATCAATTAACAGAA 332206
 Oy 238 TTTAAA-ACAAAATCATGATGATCTCATATGATGCTG-AAAAACATTTGACAAAT 295
 Db 332207 TTTAAATTTAAAATCATGATGATCTCATATGATGCAAGAAAAGCATTTTACAGAAAT 332266
 Oy 296 CTAATATTTCTTATATTAATTAATCTTCAGCAAAATCGACATGAAAGGACATACCTTA 355
 Db 332267 CAGATATCCCTTATATGATTAATAAATCATGCAAAAATGAGATGAAAGACATACCTTA 332326
 Oy 356 TGTATTAAGGACATATATATGACGAGCCACAGCAAAATTAATCTGATGGGAAAAAGTT 415
 Db 332327 GGTATTAAGGACATATATATGACGAGCCACAGCAAAATTAATCTGATGGGAAAAAGTT 332386
 Oy 416 GAAAACATTTGCTCTGAGAACTGGAACAGACAAAGATG-CTACTTTCACTCTTATTT 474
 Db 332387 GAAACATTTGCTCTGAGAACTGGAACAGACAAAGATGCTCTTCACTCTTATTT 332446
 Oy 475 CAACATATGATGAGAAAGTTTATGACGAGCAATCAACAAAGAAAGAAATCAAGGCGAC 534
 Db 332447 CAACATATGATGAGAAAGTTTATGACGAGCAATCAACAAAGAAAGAAAGGCGAT 332502
 Oy 535 CCAATATCAATTAAGAGAGATCAAACTGCTCTGTTCACTGATGATGATGATGATGATGAT 594
 Db 332503 CCAATATCAATTAAGAGAGATCAAACTGCTCTGTTCACTGATGATGATGATGATGATGAT 332562
 Oy 595 AGAAAACCTTAAGACTCATCCAGAAAGCTCTAGAACTGATATCATTAATTCAGTAAAGT 654
 Db 332563 AAAAAACCTTAAGACTCATCCAGAAAGCTCTAGAACTGATATCATTAATTCAGTAAAGT 332622
 Oy 655 TTCAAGATCAAACTTAATGTAACAATTCAGTATGATGATGATGATGATGATGATGATGAT 714
 Db 332623 TTCAAGATCAAACTTAATGTAACAATTCAGTATGATGATGATGATGATGATGATGATGAT 332682
 Oy 715 AGCTGAGATCAAACTCAAGAACTCAAACTTTTACATATAGCTGTAAT-AAAAAT 767
 Db 332683 AGCTGAGATCAAACTCAAGAACTCAAACTGCTTCACTATGCTGCAAAAGAAATTAAT 332742
 Oy 768 ACTTAAGATATTTCTTAACCAAGAGGTGAAGACTCTTACAAAGAAATCTACAAACAC 827
 Db 332743 ACTTAAGATATTTCTTAACCAAGAGGTGAAGACTCTTAAAGAAATCTACAAACAC 332802
 Oy 828 AGCT----GACATATGATGACAAACAAAGTGAAGAACTATCCCATGCTCATGATGG 883
 Db 332803 TCTGTAAGAAACAACTATGAGCAACAAACAAATGGAACCACTTCCATGCTCATGATGG 332862
 Oy 884 GTTAGATCATATATTTGTAAGAAATGACCATTTGCAAAAGCAATCTACAGTTCAATGCA 943
 Db 332863 GTTAGATCATATATTTGTAAGAAATGACCATTTGCAAAAGCAATCTACAGTTCAATGCA 332922

332863 GTTAGATCATATATTTGTAAGAAATGACCATTTGCAAAAGCAATCTACAGTTCTGCA 332922
 Oy 944 TTCCCAACCAATATCATCATCTTCAAGAACTAGAAAGAAACAAATCTTAATATTC 1003
 Db 332923 TTCCCAATTAATATCATCATCTTCAAGAACTAGAAAGAAACAAATCTTAATATTC 332976
 Oy 1004 ATATGAAACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1063
 Db 332977 ATATGAAAC-----CAAAAAGGCCCAATATGCAAAAGGCTTAAGCA 333022
 Oy 1064 AAAAGAAACAAATCTGAGGATCACTATCCCATTTCAATCTATCTACAAAGCTATTA 1123
 Db 333023 AAAAGAAACAAATTTGAGGATCACTATCCCATTTCAATCTATCTATCTAG----- 333075
 Oy 1124 TCACCAAAACATCATGAGCACTGACATTAATCTAGGCAATATGACCAATGAAAGAGAG 1183
 Db 333076 TCACCAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 333135
 Oy 1184 AGAATCCAGAAATTAAGCCAAATTAATTAAGCCAACTGATTTTGAACAAAGCAACAAA 1243
 Db 333136 AGAATCCAGAAATTAAGCCAAATTAATTAAGCCAACTGATTTGGAACAAAGCAACAAA 333195
 Oy 1244 ACATTAAGTGGGAAAGAAAGCACTTATGATTAACAAATGCTGATGATTTTGGCAAGCA 1303
 Db 333196 ACATTAAGTGGGAAAGAAAGCACTTATGATTAACAAATGCTGATGATTTGGAAGCA 333254
 Oy 1304 CATGCGAAGATTAAGCACTGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1363
 Db 333255 GGTGTAAGAAAGCTGAACTGATGATG-CTTATCTCTCACTTATACAAATTAACCTCAAG 333313
 Oy 1364 TGGATCAAGACTTAATCTGAGCTTAACCAATTAATTTAGAAATTAATCAATCAATCA 1423
 Db 333314 TGGATCAAGACTTAATCTGAGCTTAACCAATTAATTTAGAAATTAATCAATCAATCA 333373
 Oy 1424 AAAATCTTCTAGACATTTAGGCAAGCACTTATGATGATGATGATGATGATGATGATGAT 1483
 Db 333374 AAAATCTTCTAGACATTTAGGCAAGCACTTATGATGATGATGATGATGATGATGAT 333433
 Oy 1484 GCAACAAACCAAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1543
 Db 333434 GCAACAAACCAAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 333493
 Oy 1544 AAAACATCATTTAGGCAAGCAACCAACCAAGGATGAGAAATCTTCAACAAAC 1603
 Db 333494 AAAACATCATTTAGGCAAGCAACCAACCAAGGATGAGAAATCTTCAACAAAC 333553
 Oy 1604 TAAGATCTGATTAAGACTTAATCTGAGATTCACAAAGAACTCAACAAATCAAGCAAG 1663
 Db 333554 TAAGATCTGATTAAGACTTAATCTGAGATTCACAAAGAACTCAACAAATCAAGCAAG 333613
 Oy 1664 AAGAAAGCAACCAATCCCATGAAAGAGTGGCTTAAGGAC-----ATGATTAAGCAATTC 1719
 Db 333614 AAGAAAGCAACCAATCCCATGAAAGAGTGGCTTAAGGAC-----ATGATTAAGCAATTC 333672
 Oy 1720 AAAAGAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1777
 Db 333673 AAAAGAAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 333732
 Oy 1778 TAGGAAATGTAATTAATCAACCTGTAATGCGATTAACCTTACTCTGCAAGATGCTAT 1837
 Db 333733 TAGGAAATGTAATTAATCAACCTGTAATGCGATTAACCTTACTCTGCAAGATGCTAT 333792
 Oy 1838 AATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1896
 Db 333793 AATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 333852
 Oy 1897 TACACTGCTGAGGAAATGAACCTTGGCAACCACTATGAAACCAAGTGGAAATTTTC 1956
 Db 333853 TACACTGCTGAGGAAATGAACCTTGGCAACCACTATGAAACCAAGTGGAAATTTTC 333912
 Oy 1957 TTAAGAACTTAAGAAATGATGATCACTTTGATCG- 1992
 Db 333913 TTAAGAACTTAAGAAATGATGATCACTTTGATCG- 333972

Qy	2058	AAGATTAATGATGACCTTCAAGCAACTAGATGGAGACCTTATTTCTTAACT	2117
Db	62535	AAGAACAAATCAATGATTTTGGACAGACCTGATGTAATTTGAGACCAATTAATTTCTAACT	62476
Qy	2118	GGGGTAACCTAGAGATGGAATAACCAACATCAATATTTCTCACTTACAAAGTGGGGCTTAA	2177
Db	62475	GAAGTAACCTAGAGATGGAATAACCAACATCAATATTTCTCACTTAAAGTGAAGCTTAA	62416
Qy	2178	GCTGT--GAGACACAGAAAGCAT-AGAAATGATATATGAACTCTGGGGGACTTGAAGGGA	2233
Db	62415	GCTGTGAGGAGAGACGCAAAAGGCATTAAGAACATATCAATGACCTTTGGGGACTTGAAGGGA	62356
Qy	2234	AGATATGGAAGAAGAGGGAGGATTAAGAACTACACAATGGGTACAGCTGACCTGAG	2293
Db	62355	AGGATGGAGAGGTACAGGGGGTAAATTAAGAACTACACTTTGGGTGAAGTACCTGCTCAA	62286
Qy	2294	GTGATGGGTCCACCAAAATCTCAGAAATTAACCACTAAAGACTTTATCCATGAGCAAC	2353
Db	62295	CTGAGAGTGCACACCAATCTCAGAAATTCGCCCTTAATACCTTATCCATGTAACCAAC	62236
Qy	2354	ACCACTGTTTCCCAAAATCCCAATGAATAATTAATTAATGATTTAATTTC	2413
Db	62235	ACCACTGTTTCCCAAAAGCATTGAAT-AAAATTAATGAATTAATTAATCTCAATAAT	62177
Qy	2414	ACGAAATTAATAAAAGTCACCTGTGAGATTATTA	2450
Db	62176	ATTGCAATGAGTGAAGTTTATTTATTCCTGGTTTACAA	62140
RESULT 4			
US-10-087-192-418/C			
; Sequence 418, Application US/10087192			
; Publication No. US20020182586A1			
; GENERAL INFORMATION:			
; APPLICANT: Morris, David W.			
; APPLICANT: Engelhard, Eric K.			
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR			
; TITLE OF INVENTION: CANCER			
; FILE REFERENCE: 52945200012			
; CURRENT APPLICATION NUMBER: US/10/087,192			
; CURRENT FILING DATE: 2002-03-01			
; PRIOR APPLICATION NUMBER: US 09/747,377			
; PRIOR FILING DATE: 2000-12-22			
; PRIOR APPLICATION NUMBER: US 09/798,586			
; PRIOR FILING DATE: 2001-03-02			
; NUMBER OF SEQ ID NOS: 2059			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 418			
; LENGTH: 209484			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc_feature			
; LOCATION: (1)...(209484)			
; OTHER INFORMATION: n = A,T,C or G			
US-10-087-192-418			
Query Match			
Beat Local Similarity 16.2%; Score 1395; DB 13; Length 209484;			
Matches 1989; Conservative 0; Mismatches 385; Indels 173; Gaps 19;			
Qy	1	TGTATGAAGCAATGTCACCTTAATGCCAAACACAGAAAGGATTA--CAAAAAAGAAA	58
Db	93988	TCATGACGCCAGAGATGCCCTTAATACCAAAACAGAGGGATACACAAAAAGGGA	93929
Qy	59	ACTATAGACCAATACCACTGATGAATATACATGACAGAAATCCCAACAAATACAGCTA	118
Db	93928	ACTAGACCAATATCCCTGATGATGACATGATCAAAATCTTCAACAAAATACAGTGA	93869
Qy	119	ACCAATTCACACGATATTC--AAGAAGTAATCCACATTTGTCAAGTGGGTTTCATACC	176
Db	93868	ACCAATTCATATGCAATTCAAAAAAGATTAATCCATTTCAAGTGGGTTTCATACC	93809

QY 177 AGGGGTGAGATGATGTTAACTACACAGATCAATTAATGTGATACATCACTAAACGA 236
 DB 93608 AGGATATGAGATGTTAACTACATTAAGTCAATTAATGTGATACACATCACTAAATGA 93749
 QY 237 ATTAAACAAATTCATGATCATCTCAATAGATGCTGAAAAGCATTTGACAAAATC 296
 DB 93748 ATTAAACAAATTCATGATCATCTCAATAGATGATTAAGCATTAAGCATTTGACAAAATC 93689
 QY 297 TAACTTTCTTATGATTTAAACCTTCAGCAAAATTCACATAGAAAGCATACCTTAAT 356
 DB 93688 CAGATCCCTTATGATTTAAACCTTCGAAAAGCATAGAAAGCATCTTAAG 93629
 QY 357 GTATTAAGGCAATATGACGAGCCACAGCAAACTTATATCTGAATGGGAAAAGTTG 416
 DB 93628 GTATTAAGGCAATATGACGAGCCACAGCAAACTTATATCTGAATGGGAAAAGTTG 93569
 QY 417 AAAACATTTGCTGAGACCTGAAACAGACAGAGATG-CTACTTTCACCACTTCTATTC 475
 DB 93568 AAAACATTTGCTGAGACCTGAAACAGACAGAGATGCTCACTTTCACCACTTCTATTC 93509
 QY 476 AACATAGTAGGAAAGTTTACGAGCAATCAGACAGAAAGAAATCAAGGCAAC 535
 DB 93508 AACAAATGATGAAAGTTTACGAGCAATCAGACAGAAAGAAATCAAGGCAAC 93449
 QY 536 CAAATCAATTAAGGAAAGTCAACTGCTCTTCTACTGATGATGATTTACTTA 595
 DB 93448 CAAATCAATTAAGGAAAGTCAACTGCTCTTCTACTGATGATGATTTACTTA 93389
 QY 596 GAAACCTTAAGATCTACAGAAAGTCTCTAGAACATGATCACTTAATTCAGTAAAGT 655
 DB 93388 GAAACCTTAAGATCTACAGAAAGTCTCTAGAACATGATCACTTAATTCAGTAAAGT 93329
 QY 656 TCAAGATACAACTTAATGTACACAAATCAGTAGCACTGCTATACCAACAGTACCA 715
 DB 93328 TCAAGATACAACTTAATGTACACAAATCAGTAGCACTGCTATACCAACAGTACCA 93269
 QY 716 GCTGAGATCAATCAAGAACTCAACCTTTTCACTTAATGCTGTA-- 762
 DB 93268 GCTGAGATCAATCAAGAACTCAACCTTTTCACTTAATGCTGTA-- 93209
 QY 763 AAAATCTTAAGATTTCTTACCAAGAGTAGAGAGCTCTACAAAGAACTCA 822
 DB 93208 AAAATCTTAAGATTTCTTACCAAGAGTAGAGAGCTCTTACAAAGAACTCA 93149
 QY 823 AACACAGCTGACAT---CATAGATGACACAAACAGTGAACACATCCCATGCTATG 878
 DB 93148 AACACAGCTGACAT---CATAGATGACACAAACAGTGAACACATCCCATGCTATG 93089
 QY 879 GATGGGTAGATCAATTTGTGAATAATGACCAATTTGCAAAAGCAATCTACAGTTCAA 938
 DB 93088 GATGGGTAGATCAATTTGTGAATAATGACCAATTTGCAAAAGCAATCTACAGTTCAA 93029
 QY 939 TGCATTTCCCAACCAAAATATCATCATTTCTTACAGAACTAGAAAAAAACATTTCTAA 998
 DB 93028 TGCATTTCCCAACCAAAATATCAACCAATTTCTTACAGAACTAG-AAAAAAACATTTCTAA 92970
 QY 999 AATTCTATGGAACACACACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCA 1058
 DB 92969 AATTCTATGGAAC-----CAAAAGAGGCTGCTAGCAAAACCAAGCT 92924
 QY 1059 TAGCAAAAAGACAAATCTGAGGACATCACTTACCACTTTCAACTATATCTACAAAGC 1118
 DB 92923 TAGCAAAAAGACAAATCTGAGGACATCACTTACCAAGTTCAAAAGATCTATTAAGGC 92864
 QY 1119 TATATCAACCAAAACATCATGCACTGACATTAAGCACTAGCACTAGCACTAGCACTAG 1178
 DB 92863 CATATGCAACCAAAAGATGATCTAGTATTAAGCACTAGCACTAGCACTAGCACTAG 92804
 QY 1179 AAGAGAAATCCAGAAATTAAGCAAAATTAATATAGCACTAGTTTGAACAAAGCAA 1238
 DB 92803 AATAGAAATCCAGAAATTAAGCAAAATTAATATAGCACTAGCTTTGCAAAAGCAA 92744

QY 1239 CAAAAACATTAAGTGGGAAAAAGCACTTCTAGTTAAACAAATGTGCTGACATTAATGGCA 1298
 DB 92743 CAAAAACATTAAGGAGGAAAAAGCACCTTATTTCAACAAATGTGCTGGGATTAATGGCA 92684
 QY 1299 AGCCACATGAGGAAGATGAAATCGATCCCTGTCTCTCACTTAATACAAAAATTTGATA 1358
 DB 92683 AGCCACATGAGGAAGATGAAATGTGAT-CTTGCTCTCTCACTTATACAAAAATTCAACT 92625
 QY 1359 CAAGATGATCAAAAGACTTAATCTGAGACCTTAACCAATTAATTAATCTAGAAATTA 1418
 DB 92624 CAAGATGATCAAGACTTAATTAAGCTTAAGACCTTAATCAATGATGAGACCTGATAGTACA 92565
 QY 1419 TCAGAAAAATGCTTCTGACATTTCACTTGAACAAAGCTTATGAGCCAAAGAACCAAAAG 1478
 DB 92564 TTTGAAAAATCCCTTCTGACATTTGCTTATGCAAAAGCTTATGACAAAGAACCAAG 92505
 QY 1479 TAAATGCAACAAACCAAAATTAATTAATGATAGACTTAATTAATTAATTAATTAAT 1538
 DB 92504 CAATGCAACCAAAATTAATTAATGATAGATGATGATGATGATGATGATGATGATGATG 92445
 QY 1539 CAGCAAAACCAATCTTATGAGCAAGCAACCAACCAACCAACCAACCAACCAACCAAC 1598
 DB 92444 CAGTAAAGAAATTAACAGAGAG-TTTACAGCAACCAACCAACCAACCAACCAACCAAC 92387
 QY 1599 CAACTAGCATCTGATTAAGATTAATTTCCGAATTCACAGAACTCAACCAATCA 1658
 DB 92386 CAACTAGCATCTGATTAAGATTAATTTCCGAATTCACAAATCTCAAAAGTCTCAAAATCA 92327
 QY 1659 GCAAGAAAGAAAGCAAAATCCATGAAAGAGTGGGCTTAAGGACATGAATGACAAATCT 1718
 DB 92326 GCAAGAAAGAAAG---AAATTCATCAAAAGATGAGGCTTAAGGACATGAATGACAAATCT 92271
 QY 1719 CAAAGAAAGATTAACAAATG-CCAAACAAACAGAAACCAATGCTTAACATCACTAAATGA 1776
 DB 92270 CAAAGAAAGATTAACAAATGAGTCCAGACATATGAAACCAATGCTTAACATCACTAAATGA 92211
 QY 1777 TTAGGAAATGTAAATC--AACTGTAAATGCAATACCACTTCTCTGCAAGATGCT 1834
 DB 92210 TTAGGAAATGTAAATC--AACTGTAAATGCAATACCACTTCTCTGCAAGATGCT 92151
 QY 1835 CATATTTTAAATCTAAATTAATTAATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1894
 DB 92150 CATATTTTAAATCTAAATTAATTAATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 92092
 QY 1895 TTTACATGCTGTGGAATGTAAATCTTGCGCAACAATAAGAAACAGTGTGAATTT 1954
 DB 92091 TTTACATGATAGTGGAAATGTAAATCTAGTACATCACTGCAAAACAGTGTGAATTT 92032
 QY 1955 TCTTAAGAACTAAAGTA----- 1973
 DB 92031 CTTTAAAGAACTAAAGTATCTTTGATCTAGCAATTTCTAATTAATTAATTAATTAATTA 91972
 QY 1974 ----- 1973
 DB 91971 AGGAAAAAGATCTTTTACAAAAAAGATACCTGTATCATGATCTTTGTAGCAGACAAAT 91912
 QY 1974 ----GATGACCAATTTGATCCAGCAATCCCAATTAATATGATTAATTA----- 2015
 DB 91911 TCGCAATTTGCAAAATATGGAACAGCCCAATATCCATCAATCAACAGTGAATTAAT 91852
 QY 2016 TATATATTTATTTATACATGAAATACACTGAGCCATTAATAAGAAATTAATATGATGCA 2075
 DB 91851 GTGGTATTTATTTATTAATCAATGAAATATCTTATGCTTAAACAGAAACCAAAACCTTGGCA 91792
 QY 2076 TTTACAGCAATCTAGATGAAATGGAAGCTTATTTCTAATGAGGCTTATCTCAGAAATGG 2135
 DB 91791 TTTGAGCAACCTGAGATGAAATGGAAGCTTATTTCTAATGAGGCTTATCTCAGAAATGG 91732
 QY 2136 AAAACCAACATCATATTTCTCACTTAACAAGTGGGCTTAAGCTGTGAGGACAGCAAGG 2195
 DB 91731 AAAACCAACATCTATGTTCTCATTTATATGAGGCTTAATTAATGAGGCTTAAGG 91672
 QY 2196 CAT-AGATGATTAATGAATCTGAGGAACTTGAAGGAAAGATGAGAAAGGCGAGGG 2254

QY	1539	CAGCAAAAACAATCATTTAGCAGGCAAAACAACCACCGAGTGAAGAAAATCTTCA	1598
Db	92444	CAGTAAAGAAATATACGACGAGC-TTACAGACACCCACAGTGGAGGA-AAATCTTCA	92387
QY	1599	CAAACTAAGCATCTGACTAAGCACTAATATCCGAAATCCAAGAACTCAAACAATCA	1658
Db	92386	CAATCTATATCATCTGACAAAGGAATTAATCTCAAAATCTACAAAGTACTCAACAAATCA	92327
QY	1659	GCAAGAAAGAAAGCAACAATTCCTGAAAGATGGCTTAAGACATGATAGCAATTC	1718
Db	92326	GCAAGAAAAA-----AAATCTCCATCAAAAAATGAGCTTAAGCAGAAATAGCAATTC	92271
QY	1719	CAAAAGAAAGATATCAAAATGG--CCAAACAACGAAAAAAATGCTTAACATCACTAATGA	1766
Db	92270	CAAAAGAAAGATATCAAAATGGTCTCCAGCATATGAAAAATGCTCAACATCACTAATGA	92211
QY	1777	TTAGGAAATGTAAATC--AAACCTGATATGCAATACACTTAATCTCTGCAAGAAATGGT	1834
Db	92210	TCAGGAAATGCAAAATCAAAACAATGCAATGCAATACCGCTCAATCTGCAAAATGGC	92152
QY	1835	CATAATTTAAAAATCTAAAAATATATAGATGTGGTGTCTGTGGTATTAAGAACTCT	1894
Db	92150	CATATATAAAAAT--AAAAATGATATGATATTTGGTGGATGCAATGAAAAAGGAACTCT	92092
QY	1895	TTTACACTGCTGGGGAAATGTAACTTGGCAACCTATGAAAAACAGTGTGAAATTT	1954
Db	92091	TTTACACTGATAGTGGGAATGTAACTATGTAACATCACTCTGAAAAACGGTGTGGATTT	92032
QY	1955	TCCTTAAGAACTAAAGTA-----	1973
Db	92031	CCTTAAGAACTAAAGTATCTTTGATCTAGCAATTTACTAATAGGATTTACTAG	91972
QY	1974	-----	1973
Db	91971	AGAAAAAGAGTCATTATACAAAAAAGATACCTGTACATGATGTTGTAGCAGCAAT	91912
QY	1974	----GATCGACATTGTATGATCAGCAATCCCATTAATATGTATAA-----	2015
Db	91911	TCGCAATTGCAAAAATATNGAACCGACGCCCAATATCCCATCAATCAACCACTGATTAATTT	91853
QY	2016	TATATATATTTATATATACATGAAATACAACTCAGCCATAAAAAAGATTAATATGATGACA	2075
Db	91851	GTGGTATTTATATATATACATGAAATATCTAATTAGCCATAACAAGAACAAATCATGGCA	91792
QY	2076	TTGACACCAATCTAGATGAAATTTGAGAACCCCTTATTTCTAATGTGGGTAATCTCAGAAATGG	2135
Db	91791	TTGCGACCAACCTGGAATGGAATTTGAGGCTATATCTGAAATGATTAATTTCCAGAAATGG	91732
QY	2136	AAAAACAACATCATATGTTCTCAGCTTAACAAGTGGGGGCTTACCTGTGAGGACACAGAGG	2195
Db	91731	AAAAACAACATCTTATGTTCTCAATATGTGGAGCTTAACCTATGAGGGCGCTTAAGG	91672
QY	2196	CAT-AGATGATATATATGAACCTGCGGAGCTTGAAGGGGAAGATGGAAGAGGCGAGGG	2254
Db	91671	CATAGAGATATATATATATGGAACCTTGGGGGACTTGGGGGAAGAGGT--GAGGGGTGGCGAGGA	91613
QY	2255	ATTAAGAACTACCAATGGGTACAGTGTACATGCTCAAGTATG---GCTGACCAAAA	2311
Db	91612	ATTAAGAACTACCAATGGGTATACAGTATGCTGCTGAGATGATGTTATATGACCAAAA	91553
QY	2312	TCTCAGAAATTTACCACTTAAGAACTTATCCATGGAAGCAAAACCAACTGTTCCCAAAA	2371
Db	91552	TCTCAGAAATTCACCACTTAAGAAATTTATTCATGTAAACAAATGCCAGCTGTTCCCAAAA	91493
QY	2372	TCCCAATGAATAAAAATATATATAT 2398	
Db	91492	ATTATATGAATAAAAATAAAAAAT 91466	

Query Match	Best Local Similarity	Score	DB	Length
Matches 1970; Conservative	0; Mismatches	381; Indels	181; Gaps	17
US-09-933-267A-1				
ORGANISM: human				
TYPE: DNA				
LENGTH: 465237				
SEQ ID NO 1				
SOFTWARE: FastSeq for Windows Version 4.0				
NUMBER OF SEQ ID NOS: 2				
PRIOR FILING DATE: 2001-04-05				
PRIOR APPLICATION NUMBER: 09/766184				
PRIOR FILING DATE: 2001-01-24				
PRIOR APPLICATION NUMBER: 09/804076				
PRIOR FILING DATE: 2001-03-13				
PRIOR APPLICATION NUMBER: 09/826314				
PRIOR FILING DATE: 2001-04-05				
NUMBER OF SEQ ID NOS: 2				
SOFTWARE: FastSeq for Windows Version 4.0				
LENGTH: 465237				
SEQ ID NO 1				
TYPE: DNA				
ORGANISM: human				
US-09-933-267A-1				
Query Match	16.1%;	Score 1390.4;	DB 9;	Length 465237;
Best Local Similarity	77.8%;	Pred. No. 0;		
Matches 1970; Conservative	0; Mismatches	381; Indels	181; Gaps	17
US-09-933-267A-1				
ORGANISM: human				
TYPE: DNA				
LENGTH: 465237				
SEQ ID NO 1				
SOFTWARE: FastSeq for Windows Version 4.0				
NUMBER OF SEQ ID NOS: 2				
PRIOR FILING DATE: 2001-04-05				
PRIOR APPLICATION NUMBER: 09/766184				
PRIOR FILING DATE: 2001-01-24				
PRIOR APPLICATION NUMBER: 09/804076				
PRIOR FILING DATE: 2001-03-13				
PRIOR APPLICATION NUMBER: 09/826314				
PRIOR FILING DATE: 2001-04-05				
NUMBER OF SEQ ID NOS: 2				
SOFTWARE: FastSeq for Windows Version 4.0				
LENGTH: 465237				
SEQ ID NO 1				
TYPE: DNA				
ORGANISM: human				
US-09-933-267A-1				
Query Match	16.1%;	Score 1390.4;	DB 9;	Length 465237;
Best Local Similarity	77.8%;	Pred. No. 0;		
Matches 1970; Conservative	0; Mismatches	381; Indels	181; Gaps	17
US-09-933-267A-1				
ORGANISM: human				
TYPE: DNA				
LENGTH: 465237				
SEQ ID NO 1				
SOFTWARE: FastSeq for Windows Version 4.0				
NUMBER OF SEQ ID NOS: 2				
PRIOR FILING DATE: 2001-04-05				
PRIOR APPLICATION NUMBER: 09/766184				
PRIOR FILING DATE: 2001-01-24				
PRIOR APPLICATION NUMBER: 09/804076				
PRIOR FILING DATE: 2001-03-13				
PRIOR APPLICATION NUMBER: 09/826314				
PRIOR FILING DATE: 2001-04-05				
NUMBER OF SEQ ID NOS: 2				
SOFTWARE: FastSeq for Windows Version 4.0				
LENGTH: 465237				
SEQ ID NO 1				
TYPE: DNA				
ORGANISM: human				
US-09-933-267A-1				
Query Match	16.1%;	Score 1390.4;	DB 9;	Length 465237;
Best Local Similarity	77.8%;	Pred. No. 0;		
Matches 1970; Conservative	0; Mismatches	381; Indels	181; Gaps	17
US-09-933-267A-1				
ORGANISM: human				
TYPE: DNA				
LENGTH: 465237				
SEQ ID NO 1				
SOFTWARE: FastSeq for Windows Version 4.0				
NUMBER OF SEQ ID NOS: 2				
PRIOR FILING DATE: 2001-04-05				
PRIOR APPLICATION NUMBER: 09/766184				
PRIOR FILING DATE: 2001-01-24				
PRIOR APPLICATION NUMBER: 09/804076				
PRIOR FILING DATE: 2001-03-13				
PRIOR APPLICATION NUMBER: 09/826314				
PRIOR FILING DATE: 2001-04-05				
NUMBER OF SEQ ID NOS: 2				
SOFTWARE: FastSeq for Windows Version				

NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 1691139
; TYPE: DNA
; ORGANISM: Human
US-10-067-514-1

Query Match 16.0%; Score 1375.6; DB 14; Length 1691139;
Best Local Similarity 77.1%; Pred. No. 6.1e-313;
Matches 1993; Conservative 0; Mismatches 409; Indels 182; Gaps 19;

QY 1 TGTATGAAGCCAAATGCTACTTTATACCAAAACGAGAAAGATAT--CAAAAAAGAA 58
DB 1325552 TCTATGAAGCCAAATGCTACTTTATACCAAAACGAGAAAGATATCAAAAAAGAA 1325511
QY 59 ACTATGACACAGTACCACTGATGATATACATGCAAGAAATCCCAACAAATATCTAGCTA 118
DB 1325612 ACTACAGATCAATTCCTGATTAACATAGTGTAAATCTTACCAAAATCTAGCTG 1325671
QY 119 ACCAATCCAAACAGCATATCAAGAAATATCCACATTTGCAAGTGGTTTCATACCAG 178
DB 1325672 ACTGAATCCAAACATATCAAGAAATATCCACATGATCAAGTGGTTTCATACCAG 1325731
QY 179 GGGTGCAG-GATAGGTTAACTACACAAGTCAATTAATGTATCATCACTAAACGAA 237
DB 1325732 GGGTGCAGAGTGGTTTAATGTATACACAAGTCAATTAATGTATACCACTAAACGAA 1325791
QY 238 TTAATAAACAATAATCAGATGATCATCTCAATAGATGCTGAAAAAGCAATTTGACAAATCT 297
DB 1325792 TTAATAAACAATAATTCATGATCATCTCAATAGATGCAAGAAAAAGCAATTCACAAATTC 1325851
QY 298 AACATTTCTTTATGATTTAAACCTTCAGCAAAATCGACATAGAAAGACATACCTTAATG 357
DB 1325852 AGCATCCCTTTATGATTTAAAGCTCTCAGCAAAATCAGCATACAGGACATACATTAATG 1325911
QY 358 TAATAAAGCATATATAGCGAGCCCAAGCAACATTAATCTGAATGGGAAAAAGTTGA 417
DB 1325912 TAATAAAGCATATCTATAGCAAAACCAAGCAACGTAATCTGAATGGGAAAAAGTTGA 1325971
QY 418 AACATTTGCTCTGAGATGGAACAGACAAAGATG-CTACTTTCAACCTTCTATTTC 476
DB 1325972 AAGATTTCCCTCTGAGATGGAACAGACAAAGATGCTCTCACTCTTCTTTCA 1326031
QY 477 ACATAGTAGTGAAGTTTAAAGCCAGACATCAGACAAAGAAAGAAATCAAGGCAACC 536
DB 1326032 ACATAGTAGTGAAGTCTCTAGCAAGACCAATCAGACAAAGAGGAAATTAAGGCAATCC 1326091
QY 537 AAATCAATTAAGGGAAGTCAAACTGCTCTGTTCACTGATGATGATGTTGATATACCTAG 596
DB 1326092 AAATCGGTAAAGGAAGTCAAACTGCTCACTGTTGCTGATGATGATGATTAATACCTTG 1326151
QY 597 AAAACCTTAAGAGCTCATCCAGAAAGCTCTGAGATGATCATTAATTCAGTAAAGTTT 656
DB 1326152 AAAACCTTAAGAGCTCTCCAGCAAGCTCTTGAACGTAAATGAAATTCAGAAAGTTT 1326211
QY 657 CAGATATCAAACTTAATGATCACAAAATCAGTGAAGCTGTATACCAACAGTGAACAG 716
DB 1326212 CTGGAATCAAGATTAATGATCACAAAATCAGTGAAGCTCTTCTATATACCAACAGTGAACAG 1326271
QY 717 CTGAGATCAAACTTAATGATCACAAACTTCTTAATTAAGCTGTAA-----AAA 765
DB 1326272 GGGAGATCAAACTTAATGATCACAAACTTCTTAATTAAGCTGTAA-----AAA 1326331
QY 766 ATACTTAAGATATTTCTTACCAAGAGTGAAGAGCTCTTACAAAGAAAACTCAAAAC 825
DB 1326332 ATACTTAAGATATTTCTTACCAAGAGTGAAGAGCTCTTACAAAGAAAACTCAAAAC 1326391
QY 826 ACAGCTGA-----CATCATAGATGACAAACAAAGTGAACAACTCCCATGCTCATGAT 881
DB 1326392 ACTGCTGAAGAGATCATAGCAATTAACAAACAAATGGGAAACACTCCCGATCATGAT 1326451
QY 882 GGGTAGAATCAATTTGTGAAGAAATGACCATATTTGCCAAAGCAATCTACAGTTCAATGC 941

DB 1326452 GGGTAGAATCAATTTGTGAAGAAATGACCATATTTGCCAAAGCAATCTACAAATTCACAC 1326511
QY 942 AATTCACCAAAATATCATCATCTTCTTACAGAACTGAAAAAACAATTTCAAAAT 1001
DB 1326512 AATCCCATCAAAATATACCATCATCTTCTTACAGAAATTTG-AAAAACAATTTCAAAAC 1326570
QY 1002 TCATATGGAACACACACCAAAAAA-----CCCCATAGCCCAAGCAAGCTTAG 1061
DB 1326572 TCATATGGAAC-----CAAAAAAGCTGCAATAGCCCAAGCAAGCAAG 1326616
QY 1062 CAAAAAGCAAAATCTGAGGCAATCATATTAACCATCTTCAAAATATATCTCAAGGCTAT 1121
DB 1326617 CAAAAAGTCAAAATCTGAGGCAATCATATTAACCATCTTCAAAATATATTAAGGCTAT 1326676
QY 1122 AATCACCAAAAATCATGAGCACTGACATTAATACTGAGCACATAGACCAATGAAAAAGAG 1181
DB 1326677 AGTCAACAAAATGAGCAATGATGATGATTAATAAATAGCATATAGACCAATGGAACAGAT 1326736
QY 1182 AGAGATTCAGAAATTAAGCCAAATTAATTAAGCCAACTGATTTTGAACAAAGCAACAA 1241
DB 1326737 AGAGAACCCGAGATTAACCCAAATTAATTAAGCCAACTGATTTTGAACAAAGCAACAA 1326796
QY 1242 AAACATTAAGTGGGGA---GACATTCATTAACAAATGCTGAGATTAATTTGGAAG 1300
DB 1326797 AAACATTAAGTGGGGAAGATTAACCTTTCAACAAATGCTGAGATTAATTTGGAAG 1326856
QY 1301 CCAATGTGAGAAAGATGAACCTGATCCCTGCTCTCACTTAATACAAAAATTTGATACA 1360
DB 1326857 CCAACATAGGAAGATGAACCTGATCC--TATCTCTCACCGTATACAAAAATCACTCA 1326914
QY 1361 AGATGATCAAGACTTAATCTGAGACTTAACCAATTAATAATTTGAGAAATACATC 1420
DB 1326914 AGATGATTAAGGCTTAACCTTAAGAGCTGAACCTAT--GAAATTTAGAAATTAACCTTT 1326973
QY 1421 AGAAAAATGCTTCAAGCAATTCCTTAGGCAAGACTTCATGCGCAAGAACCCAAAGTA 1480
DB 1326973 GAAAAATCCCTTCAAGCAATTCCTTAGGCAAGACTTCATGCGCAAGAACCCAAAGTA 1327033
QY 1481 AATGCAACAAAAAATAAATTAATAGATAGCACTTAATTAATTAAC-TAAAAAGCTTTGCGC 1539
DB 1327034 AATGCAATTAATAAATAAATTAATAGATAGCACTTAATTAATTAATTAATTAATTAATTA 1327093
QY 1540 AGCAAAACCAATCAATTAAGCAAGCAACCAACCAAGTGAAG-AAAACTTTCA 1598
DB 1327094 GGCAAAAGCAAGTGAAGCAAGTGAACCAACCAAGTGAAG-AAAACTTTCA 1327153
QY 1599 CAATCTAAGCATCTGATTAAGCACTAATCCGGAATCCCAAGGAATCAAAATCA 1658
DB 1327154 CAATCTAAGCATCTGATTAAGCAAG-GTAGTATCCAGAAATCTCAAGGACCCCAAAATCA 1327212
QY 1659 GCAAGAAAGAAAGCAACATCCCATGAAGAGTGGCTTAAGGACATGAATAGCAATTCCT 1718
DB 1327213 GTTAAGAAAAAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1327272
QY 1719 CAAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1776
DB 1327273 CAAAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1327332
QY 1777 TTAGGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1834
DB 1327333 TCAAGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1327392
QY 1835 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1892
DB 1327393 CATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1327452
QY 1893 CTTTACACTGCTGCTGGAATGTAATTTGCGCAACCACTATGGAACCAAGTGTGA 1952
DB 1327453 CTTTACACTGCTGCTGGAATGTAATTTGCGCAACCACTATGGAACCAAGTGTGTGA 1327512
QY 1953 TTTCTTAAGACCTAAAGATGATTCGACATTTGATCAGCAATTCAT----- 2001

QY	2002	-----	-----	2001
Db	1327513	TTCCTTAAAGAACTAAAAGTAGAACTACACCTTGATCCAGACGTCCTACTAGATATC	1327512	
QY	2002	-----	-----	2001
Db	1327573	TACCAGAGAAAAAGAGTCATTATTGTAAAAAAGACCTGTGACGTATGTTATAGCA	1327632	
QY	2002	-----	-----	2001
Db	1327633	GCACATTCCAAATTGCCAAAACCTGTGGAACTTACCCAAATGTCATCACTAGTCAATTACGTG	1327692	
QY	2002	-TAAATATGTATAAATATATATATATTTATATACATGAAATACAACTCAGCCATAAAAAG	2060	
Db	1327693	ATAAAGAACTGTGTGTACAGATATATATCAATGAATATCTACAGCTATGAAAAAG	1327752	
QY	2061	AATAAAATGATGACACTTACACGCAATCTAGATGGAATTGGAGACCCCTTATTTCTAAGTGG	2120	
Db	1327753	AATGAAATTACAGCACTTTGCACTGTGACCTGGATGGAAATGGAGACTATATTCTAAGTGAA	1327812	
QY	2121	GTAATCTCAGGAATGGAAAAACCAACATCATATGTTCTCATCTACAGTGGGGGCTAAGCT	2180	
Db	1327813	GTAATCTCAGGAATGGAAAAACCAACATCATATGTTCTCATCTGATATGTTGGGAATCTAAGCT	1327872	
QY	2181	GTGAGACACGAAAGGCAT-AGATGATATTAATGAACCTCTGGGGACTTGAGGGGAAGATG	2239	
Db	1327873	ATGAGGACACAAAGATATATAGAAATATATCAATGACCTTTGGGGACTTGAGGGGAAGAGTG	1327932	
QY	2240	GAAAGAGCCGAGGATATAAGACTACACATGCGTACAGTGTACACTGCTCAGGTGATG	2299	
Db	1327933	GGAGGGGGGTGAGGATTAAGATTACAAATATGTGTGACGTGTATACCTGTGGGTGATG	1327992	
QY	2300	GGTGACCAAAATCTCAGAAATATACCATTAAGAACTTATCATGAGGAACAA--ACACCA	2357	
Db	1327993	GGTGACCAAAATCTCAGAAATATACCATTAAGAACTTATCATGAGGAACCTTATTAATTAACCA	1328052	
QY	2358	CCTGTTCCCAAAATCCCAATGAATAAAAATATATATATTAATGATTTAATTTTCCAG	2417	
Db	1328053	CCTGATCCCTATATACTATTATGAAAAAAATTTTATAAAAAGTAATGATTTAAGTCAGAA	1328112	
QY	2418	AATT 2421		
Db	1328113	GTTT 1328116		
RESULT 8				
US-10-419-723-1				
; Sequence 1, Application US/10419723				
; Publication No. US20040014099A1				
; GENERAL INFORMATION:				
; APPLICANT: Greteardottir, Solveig				
; APPLICANT: Jonsdottir, Sif				
; APPLICANT: Reynisdottir, Sigridur Th.				
; APPLICANT: Thorleifsson, Gudmar				
; APPLICANT: Gulcher, Jeffrey				
; TITLE OF INVENTION: SUSCEPTIBILITY GENE FOR HUMAN STROKE;				
; TITLE OF INVENTION: METHODS OF TREATMENT				
; FILE REFERENCE: 2345.2010-005				
; CURRENT APPLICATION NUMBER: US/10/419,723				
; CURRENT FILING DATE: 2003-04-18				
; PRIOR APPLICATION NUMBER: 10/255,120				
; PRIOR FILING DATE: 2002-09-25				
; PRIOR APPLICATION NUMBER: 10/067,514				
; PRIOR FILING DATE: 2002-02-04				
; PRIOR APPLICATION NUMBER: 09/811,352				
; PRIOR FILING DATE: 2001-03-19				
; NUMBER OF SEQ ID NOS: 87				
; SOFTWARE: FastSeq for Windows Version 4.0				
; SEQ ID NO 1				
; LENGTH: 1691139				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
US-10-419-723-1				

Best Local Similarity 77.1%; Pred. No. 6.1e-313;
Matches 1993; Conservative 0; Mismatches 409; Indels 182; Gaps 19;

QY	1	1GTAAGAACCCAAATGTCATCTTTAAATACCAAAACAGGAAAGAGATAT-CAAATAAATA	58
Db	132552	TCATAGAACCCAGCATCACCTTAGTACCAAAACCGAAGAGACATACCAAAATAAATA	1325611
QY	59	ACTATGACCAAGTACCACTGATGATATATACGAGAAATCCCAACAAATACTAGCTA	118
Db	1325612	ACTACAGATCAATATCTTTGATTAACATAGATGTTAAATCTTTAAACAAATACTAGCTG	1325671
QY	119	AACCAATCCACAGATATCAAGAAAGATATCCACCAATGTCAGTGGGTTCTTACAG	178
Db	1325672	ACTGATCCAAACACATATCAAAAAGATATCCACCATGATCAAGTGGGTTCTTACAG	1325731
QY	179	GGGTGACG-GATAGGTTAAACATACACAGATCAATAAATGTGATATCATCATAAACGAA	237
Db	1325732	GGGTGACAGATGTTTAAATGTATACCAAGTCAATAAATGTGATATCACACATAACGAA	1325791
QY	238	TTAAAACAAAATCACATGATCATCTCAATAGATGCTGAAAAAGCATTTGACAAATCT	297
Db	1325792	TTAAAACAAAATTCATGATCATCTCAATAGATGACGAAAAAGCATTCACAAATTC	1325851
QY	298	AAATTTCTTTATGATTTAAACCTTCAGCAAAATTCACATGAAAGACATACCTTATG	357
Db	1325852	AGCATCCCTTTATGATTTAAACCTTCAGCAAAATTCAGATACAGGACATACATTAATG	1325911
QY	358	TAAATAAAGCATATATAGACGAGCCACAGCAAACTTTATCTGATATGGGAAAAAGTTGA	417
Db	1325912	TAAATAAACATATCTATATACAAACCCACAGCCACATGTAATATGAGGAAAAAGTTGA	1325971
QY	418	AAACATGTCCTCTGAGAACTGGAACAGACAGAGATG-TACCTTACACACTCTTATTC	476
Db	1325972	AAGATTCCTCTGAGAACTGGAACAGACAGATATGCCCCTCTCACACTCTTCTTCA	1326031
QY	477	ACATGTATGTGAAGTTTATAGCCAGAGAACTACAGAAAGAAATCAAGGGACCC	536
Db	1326032	ACATGTATGTGAAGTCTTACAGAGAACTACAGAAAGAAATCAAGGGACATCC	1326091
QY	537	AAATCAATAAAGAGAGATCAAACTGTCTCTGTTCATGATGATATGTTGTATACCTAG	596
Db	1326092	AAATGGTAAAGAGAGATCAAACTGTCTCTGTTCATGATGATATGTTGTATACCTAG	1326151
QY	597	AAAACTTAAGACTCATCCAGAAAGCTCCTAGAACCTGATCATAAATTCATGAAAGTTT	656
Db	1326152	AAAACTTAAGACTCCTCCAGAAAGCTCCTAGAACCTGATTAAGAAATTCAGAAAGTTT	1326211
QY	657	CAGATTAACAACTAAATGTACAAATAGTAGACTGTATATACCAACAGTACCAAG	716
Db	1326212	CTGATTAACAAAGTATATGTATACAAATAGTAGCTCTTATATATACCAACAGTACCAAG	1326271
QY	717	CTGAGATCAAAATCAAGAACTCAAAACATTTTACATAGCTGTAA-----AAA	765
Db	1326272	GGGAAATCAAAATCAAGAACTCAAAACATTTTACATAGCTGTAAATAAATAAATAAATA	1326331
QY	766	ATACTTAAGATATTTCTTACCAGAGAGGTAAAGAGCTCTACAGAGAAACCTACAAAC	825
Db	1326332	ATACTTAAGATATTTCTTACCAGAGAGGTAAAGAGACTTCTACAGAGAAACCTACAAAC	1326391
QY	826	ACAGCTGA-----CATCATAGATGACAGAAAGAGAAACACATCCCATGCTCATGAT	881
Db	1326392	ACTGTGAAGAGAAATCATGACATATCAACAAATAATGGAAACACATCCCGTATATGAT	1326451
QY	882	GGGTGATCAATATTTGTAAATATGACCATATTTGCCAAAGCAATCTTACAGTTCATATG	941
Db	1326452	GGGTGATCAATATTTGTAAATATGACCATATTTGCCCATATGCTACCAATTCATCAAC	1326511
QY	942	AATTCACACAAATATATCATCATATCTTCTACAGAACTAGAAAAAAACAATTTTAAAT	1001
Db	1326512	AATTCACACAAATATATCAACATCATATCTTCTACAGAAATTTAG-AAAAACAATTTTAAAC	1326570
QY	1002	TCATATGGAACAACAACCAAAAAAAAAAAAAAAAACCCGCATATACCAAGAACATTTAG	1061

QY 7196 TCAGGCCAGATAGATATCATCGGGTCCACAGCACTGGCTAGATGAGTGGGGGTT 7255
 Db 1 TCAGGCCAGATAGATATCATCGGGTCCACAGCACTGGCTAGATGAGTGGGGGTT 60
 QY 7256 TTGATCTTAATGTTATCCCATGTCAGACAGAACTTGTGTGCAAGAGAGGTGAG 7315
 Db 61 TTGATCTTAATGTTATCCCATGTCAGACAGAACTTGTGTGCAAGAGAGGTGAG 120
 QY 7316 GCTTCAGAGTCAACAAGAACTGATTTTCAAACTGAGATTGAGAACCCCACTTTTGATA 7375
 Db 121 GCTTCAGAGTCAACAAGAACTGATTTTCAAACTGAGATTGAGAACCCCACTTTTGATA 180
 QY 7376 GGTGACTTAATCTCTGAGATCTCTGATCTCTCTCTTTAATGAGACAGTAATCCCA 7435
 Db 181 GGTGACTTAATCTCTGAGATCTCTGATCTCTCTCTTTAATGAGAGAACTTAATCCCA 239
 QY 7436 CATGGCAGGGTGGTGGGAGAAATCAGAGATCAAAACAGCTGGTATCAATCTGGTTCTG 7495
 Db 240 CATGGCAGGGTGGTGGGAGAAATCAGAGATCAATACAGCTGGTATCAATCTGGTTCTG 239
 QY 7496 TTTCCAGAGGTCAACAGCTGGGTTCTGAGCAATGATTTCAACATCCCACTTGGGTA 7555
 Db 300 TTTCCAGAGGTCAACAGCTGGGTTCTGAGCAATGATTTCAACATCCCACTTGGGTA 359
 QY 7556 CAGAACTGACCAATCAACGAGAGTGAAGAGACTCTCTCTCAAGCAGACCTTGAGCT 7615
 Db 360 CAGAACTGACCAATCAACGAGAGTGAAGAGACTCTCTCTCAAGCAGACCTTGAGCT 419
 QY 7616 TCACGAGGCTGACAGTCACTGTTTCCCTTGTGCGCTGACAGAAACGCGTTGTGCT 7675
 Db 420 TCACGAGGCTGACAGTCACTGTTTCCCTTGTGCGCTGACAGAAACGCGTTGTGCT 479
 QY 7676 GGCCTCTGGGGTGGCGGATGCGGAGAAAGCTGTCTCACTCACTCACTCACTCACTGCTG 7735
 Db 480 GGCCTCTGGGGTGGCGGATGCGGAGAAAGCTGTCTCACTCACTCACTCACTCACTGCTG 539
 QY 7736 CGGCGGACTCTCTCTCTTCTAGCGGACATTAATGTTGCGCGTTAGCGCTCATCATTA 7795
 Db 540 CGGCGGACTCTCTCTCTTCTTCTAGCGGACATTAATGTTGCGCGTTAGCGCTCATCATTA 599
 QY 7796 TCCGCGATCCCATCTCAAAATCTCTCACTCTGTGATGACCTTTCCTTCACTTAATGAGCC 7855
 Db 600 TCCGCGATCCCATCTCAAAATCTCTCACTCTGTGATGACCTTTCCTTCACTTAATGAGCC 659
 QY 7856 TAAGATGCTGAGCGGATCAACGAGAGCTGCTCTCACTCTGTGAGCTGCTGCTGCTGCT 7915
 Db 660 TAAGATGCTGAGCGGATCAACGAGAGCTGCTCTCACTCTGTGAGCTGCTGCTGCTGCTGCT 719
 QY 7916 ACCACTGCGCGCGCCAGATACCTGTCACTGGTCACTGTCTGTCTGTCTGTGGGCGCTGT 7975
 Db 720 ACCACTGCGCGCGCCAGATACCTGTCACTGGTCACTGTCTGTCTGTCTGTGGGCGCTGT 779
 QY 7976 CCTGTGCGGAGATATCTGTGAGATGATTTCTGTGACCTTCTGTGTTTATGTTGTTGTT 8035
 Db 780 CCTGTGCGGAGATATCTGTGAGATGATTTCTGTGACCTTCTGTGTTTATGTTGTTGTT 839
 QY 8036 CTGTTTGTGTAAGATGATGATTTTCAATTAACATGCGGTGGCTGTTTATGTTGTTGTT 8095
 Db 840 CTGTTTGTGTAAGATGATGATTTTCAATTAACATGCGGTGGCTGTTTATGTTGTTGTT 899
 QY 8096 TTCTCTGTGGGTTCACAGCTGTCTGTCTGTGAGATTTCTGTGATTTCCCGAAGATGTC 8155
 Db 900 TTCTCTGTGGGTTCACAGCTGTCTGTCTGTGAGATTTCTGTGATTTCCCGAAGATGTC 959
 QY 8156 CGCTGACAGGCTGTAAGTACATCTCTCTCAAGTGTGTTCTCTCTCTGTGGCC 8215
 Db 960 CGCTGACAGGCTGTAAGTACATCTCTCTCAAGTGTGTTCTCTCTCTGTGGCC 1019
 QY 8216 TGCCCTTTGGCAATTCAGTGGGCGCTTTTCCAGAGATCCCACTGGATTTGAAAGTCTTAT 8275
 Db 1020 TGCCCTTTGGCAATTCAGTGGGCGCTTTTCCAGAGATCCCACTGGATTTGAAAGTCTTAT 1079
 QY 8276 TTTGTATGTGATCTAGATTTTCAATTTCTGTCTGCTCTTAACAGCAGTGCACCA 8335

Db 1080 TTTGTATGTGATCTAGATTTTCAATTTCTCTGTCCGCTCTTAACAGCAGTGCACCA 1139
 QY 8336 TCATTTACTTTCTCTGCGGCTCTTTTACGACAGCTCAAAATAGCAGAACTGAAAGCTG 8395
 Db 1140 TCATTTACTTTCTCTGCGGCTCTTTTACGACAGCTCAAAATAGCAGAACTGAAAGCTG 1199
 QY 8396 TTCTCAGAGGGCTCTGACAGACAGCTGAGAGTGAATGAAGTGAAGGTTGCTTCTC 8455
 Db 1200 TTCTCAGAGGGCTCTGACAGACAGCTGAGAGTGAATGAAGTGAAGGTTGCTTCTC 1259
 QY 8456 AGGAAACCTTGAGAGCTGTGCGGAGAGAGATTTGAGAGAGTGAAGAACTTGTGCTG 8515
 Db 1260 AGGAAACCTTGAGAGCTGTGCGGAGAGAGATTTGAGAGAGTGAAGAACTTGTGCTG 1319
 QY 8516 CAGACAGACTTTGAGAGCAATGCTGCCCTGACACCTTGAACATTAATGATTTTCT 8575
 Db 1320 CAGACAGACTTTGAGAGCAATGCTGCCCTGACACCTTGAACATTAATGATTTTCT 1379
 QY 8576 TAGCTTCTGCTCAGAAATG 8596
 Db 1380 TAGCTTCTGCTCAGAAATG 1400

RESULT 10

US-10-225-567A-673
 ; Sequence 673, Application US/10225567A
 ; Publication No. US20030113798A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lifespan Biosciences
 ; APPLICANT: Brown, Joseph P.
 ; APPLICANT: Burner, Glenn C.
 ; APPLICANT: Roubh, Christine L.
 ; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
 ; FILE REFERENCE: 1920-4-4
 ; CURRENT APPLICATION NUMBER: US/10/225,567A
 ; PRIOR FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 60/257,144
 ; NUMBER OF SEQ ID NOS: 2292
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 673
 ; LENGTH: 1400
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-225-567A-673

Query Match 15.9%; Score 1374.6; DB 15; Length 1400;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 1391; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

QY 7196 TCAGGCCAGATAGATATCATCGGGTCCACAGCACTGGCTAGATGAGTGGGGGTT 7255
 Db 1 TCAGGCCAGATAGATATCATCGGGTCCACAGCACTGGCTAGATGAGTGGGGGTT 60
 QY 7256 TTGATCTTAATGTTATCCCATGTCAGACAGAACTTGTGTGCAAGAGAGGTGAG 7315
 Db 61 TTGATCTTAATGTTATCCCATGTCAGACAGAACTTGTGTGCAAGAGAGGTGAG 120
 QY 7316 GCTTCAGAGTCAACAAGAACTGATTTTCAAACTGAGATTGAGAACCCCACTTTTGATA 7375
 Db 121 GCTTCAGAGTCAACAAGAACTGATTTTCAAACTGAGATTGAGAACCCCACTTTTGATA 180
 QY 7376 GGTGACTTAATCTCTGAGATCTCTGATCTCTCTCTTTAATGAGACAGTAATCCCA 7435
 Db 181 GGTGACTTAATCTCTGAGATCTCTGATCTCTCTCTTTAATGAGAGAACTTAATCCCA 239
 QY 7436 CATGGCAGGGTGGTGGGAGAAATCAGAGATCAAAACAGCTGGTATCAATCTGGTTCTG 7495
 Db 240 CATGGCAGGGTGGTGGGAGAAATCAGAGATCAATACAGCTGGTATCAATCTGGTTCTG 299
 QY 7496 TTTCCAGAGGTCAACAGCTGGGTTCTGAGCAATGATTTCAACATCCCACTTGGGTA 7555

Db 300 TTTCAGGGGTCACGAGCTGGGGTTTCTGAGCATGATTCACCATCCAGTCTTGGGTA 359
 QY 7556 CAGAACTGACACCAATCAACGAGAGTGAAGACTCTTGTCTACAGACAGACCTTGAGCT 7615
 Db 360 CAGAACTGACACCAATCAACGAGAGTGAAGACTCTTGTCTACAGACAGACCTTGAGCT 419
 QY 7616 TCACAGGGGGCTGACCTGATCGTTTCCCTTGTGCGGCTGACAGGAAAGCGGGTGTGCTCT 7675
 Db 420 TCACAGGGGGCTGACCTGATCGTTTCCCTTGTGCGGCTGACAGGAAAGCGGGTGTGCTCT 479
 QY 7676 GGCTCTCGGGCTGCGGATGCGAGAAAGCTGTCTCATATGATCTCAACCTGGATG 7735
 Db 480 GGCTCTCGGGCTGCGGATGCGAGAAAGCTGTCTCATATGATCTCAACCTGGATG 539
 QY 7736 CCGGCGGACTTCTCTTCTTACGCGGCAATTATATGTTCGCGCTTACGCTCATCAATA 7795
 Db 540 CCGGCGGACTTCTCTTCTTACGCGGCAATTATATGTTCGCGCTTACGCTCATCAATA 599
 QY 7796 TCCGCCATCCCATCTCCAAAAATCTCAAGTCTGTGATGAACCTTCCCTTATAGGCC 7855
 Db 600 TCCGCCATCCCATCTCCAAAAATCTCAAGTCTGTGATGAACCTTCCCTTATAGGCC 659
 QY 7856 TTAACATGCTGAGGCGCATGACGACGAGGCGTGTGATGCTGTGCGCCATCTGT 7915
 Db 660 TTAACATGCTGAGGCGCATGACGACGAGGCGTGTGATGCTGTGCGCCATCTGT 719
 QY 7916 ACCATGCGCGCGCCGAGATACCTGATCATGATGATGCTGTGCGCCATCTGT 7975
 Db 720 ACCATGCGCGCGCCGAGATACCTGATCATGATGATGCTGTGCGCCATCTGT 779
 QY 7976 CCGTGTGCGGAGATATCTGAGATGAGATGTTCTGTGATCTCTGTTAGTGTGCTGAT 8035
 Db 780 CCGTGTGCGGAGATATCTGAGATGAGATGTTCTGTGATCTCTGTTAGTGTGCTGAT 839
 QY 8036 CTGTTGTGTGAACGCTGAGATTTCTTACATCGGCTGCTGTTTATAGTGG 8095
 Db 840 CTGTTGTGTGAACGCTGAGATTTCTTACATCGGCTGCTGTTTATAGTGG 899
 QY 8096 TTCTGTGTGGTCCAGCTGTGCTGTGATCAGATTTCTGTGATTCGCGAAGATGC 8155
 Db 900 TTCTGTGTGGTCCAGCTGTGCTGTGATCAGATTTCTGTGATTCGCGAAGATGC 959
 QY 8156 CCGTGTGCGGAGATATCTGAGATGAGATGTTCTGTGATCTCTGTTAGTGTGCTGAT 8215
 Db 960 CCGTGTGCGGAGATATCTGAGATGAGATGTTCTGTGATCTCTGTTAGTGTGCTGAT 1019
 QY 8216 TGCCCTTTGGCATTCAGTGGGCGCTGTTTCCAGATCCACCTGATTTGAAAGCTTAT 8275
 Db 1020 TGCCCTTTGGCATTCAGTGGGCGCTGTTTCCAGATCCACCTGATTTGAAAGCTTAT 1079
 QY 8276 TTGTGATGTGATCTAGATTTCAATTTTCTGTGCTGCTTAAAGAGTGCACCCCA 8335
 Db 1080 TTGTGATGTGATCTAGATTTCAATTTTCTGTGCTGCTTAAAGAGTGCACCCCA 1139
 QY 8336 TCATTTACTTCTTGTGTGGGCTCTTTAGGAGCGGTCAAAATAGGAGAACTTGAAGCTG 8395
 Db 1140 TCATTTACTTCTTGTGTGGGCTCTTTAGGAGCGGTCAAAATAGGAGAACTTGAAGCTG 1199
 QY 8396 TTCTCCAGAGGCTCTGAGAGGACGCTGATGATGATGAGAGGAGGCTTCTC 8455
 Db 1200 TTCTCCAGAGGCTCTGAGAGGACGCTGATGATGATGAGAGGAGGCTTCTC 1259
 QY 8456 AGGAAACCTTGAGCTGTGCGGAGAGAGATTTGAGAGAGAGAACTTCTCCTGT 8515
 Db 1260 AGGAAACCTTGAGCTGTGCGGAGAGAGATTTGAGAGAGAGAACTTCTCCTGT 1319
 QY 8516 CAGACAGAGACTTTAGAGCAATGCTGCGCTGCGACCTTGAATATATGATTTTCT 8575
 Db 1320 CAGACAGAGACTTTAGAGCAATGCTGCGCTGCGACCTTGAATATATGATTTTCT 1379
 QY 8576 TAGCCTTCTGCTCAGAAATG 8596
 Db 1380 TAGCCTTCTGCTCAGAAATG 1400

RESULT 11
 US-10-322-696-4/c
 ; Sequence 4, Application US/10322696
 ; Publication No. US20040166490A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Malandro, Marc
 ; TITLE OF INVENTION: NOVEL THERAPEUTIC TARGETS IN CANCER
 ; FILE REFERENCE: 529452001200
 ; CURRENT APPLICATION NUMBER: US/10/322,696
 ; CURRENT FILING DATE: 2003-10-17
 ; NUMBER OF SEQ ID NOS: 186
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 100944
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-322-696-4

Query Match 15.9%; Score 1370; DB 17; Length 100944;

Best Local Similarity 78.1%; Pred. No. 2,2e-312; Mismatches 355; Indels 210; Gaps 22;

Matches 2013; Conservative 0; Mismatches 355; Indels 210; Gaps 22;

QY 1 TGTATGAAGCCAAATGTGACTTTAATACCAAAACGAGAAAGATATA--CAAAAAGAAA 58
 Db 46567 TGTATGAAGCCAAATGTGACTTTAATACCAAAACGAGAAAGATATA--CAAAAAGAAA 46508
 QY 59 ACTATGACCAATGACCTGATGATATATACATGAGAAATCCCAACAAATATAGCTA 118
 Db 46507 ACTATGACCAATGACCTGATGATATATACATGAGAAATCCCAACAAATATAGCTA 46448
 QY 119 ACCAATCCAAACGATATCAAGAGATATCCACATTTGCAAGTGGTTTCATACAG 178
 Db 46447 ACCAATCCAAACGATATCAAGAGATATCCACATTTGCAAGTGGTTTCATACAG 46388
 QY 179 GGGTGAGG--ATAGTTAATACATCAAGTCAATTAATGTGATATCATCATATAACAGAA 237
 Db 46387 GGGTGAGG--ATAGTTAATACATCAAGTCAATTAATGTGATATCATCATATAACAGAA 46328
 QY 238 TTAATAACAAATCAATGATATCTCATATGATGCTGAAAAAGATTTGACAAATCT 297
 Db 46327 TTAATAAC--AAATCAAGATATCTTATGATGATGAGAAAAAGATTTGACAAATCT 46269
 QY 298 AACATTTCTTATGATTAATAACCTTCAGCAAAATCGACATAGAAAGACATACCTAATG 357
 Db 46268 AACATTTCTTATGATTAATAACCTTCAGCAAAATCGACATAGAAAGACATACCTAATG 46210
 QY 358 TAATTAAGCCATATATATGAGCGAACCAAGCAAAATATATATGAGGAAAAAGTTGA 417
 Db 46209 TAATTAAGCCATATATATGAGCGAACCAAGCGCTAATATATATGAGGAAAAAGTTGA 46150
 QY 418 AAACATTTCTTATGATTAATAACCTTCAGCAAAATCGACATAGAAAGACATACCTAATG 477
 Db 46149 AAACAT--TCCCTGAGAAACCGGAAACAGAGATGCACTTACACACTTATTCAT 46092
 QY 478 CATAGTAGGAAGTTTATGAGAGCAATCAGACAGAGAAAGAAATCAAGGGCACCAC 537
 Db 46091 CATAGTAGGAAGTTTATGAGAGCAATCAGACAGAGAAAGAAATCAAGGGCACCAC 46032
 QY 538 AATCAATTAAGAGAGAACTGATCTGCTGTTCACTGATGATATGATTTATATAGTA 597
 Db 46031 AATCAATTAAGAGAGAACTGATCTGCTGTTCACTGATGATATGATTTATATAGTA 45975
 QY 598 AAACCTTAAGAGCTATCCAGAAAGCTCTTGAAGTATATATATATATATATATATAT 657
 Db 45974 AAACCTTAAGAGCTATCCAGAAAGCTCTTGAAGTATATATATATATATATATATAT 45915
 QY 658 AGATACAAACTAATATGATACAAATCAGTACGCTGATATACCAAGAGTCAAGC 717
 Db 45914 AGATACAA--ATATATATGATAGCTGCTGATATACCAAGAGTCAAGC 45865

QY 2031 ACCATGATATACACTCAGCCATTAATAAAGATATATATGATGATTCACAGCAATCTAG 2090
 Db 220647 ACAATGAAATTTTACTCGGCATTTAAAGGATTAATTAATGATTCATTTGACGACCCGG 220706
 QY 2091 ATGGAATGAGAACCTTATTTCTAAGTGGGGTAACTCAGGAATGAGAAACCAACATCAT 2150
 Db 220707 ATGACATGGAACATTTATTTCTAAGTGAAGTAACAGGAATGAGAAACCAACATTTG 220766
 QY 2151 ATGTTCTCATTAACAAGTGGGGCTAAGCTGTGAGACACGAAGGCAT -AGAATGATATA 2209
 Db 220767 ATGCTCTCCCTGATATGTGGAGCTAAGCTATGAGACGCAAGGATTAAGATGATAC 220826
 QY 2210 ATGAATCTGGGAGCTTGAAGGGGAGATGAGAAAGGCGAGGATTAAGATGATAC 2269
 Db 220827 AGGACCTTTGGGAGCTTGGGGAGAGGTGGAGTGGGCGAAGGATTAAGATGATAC 220886
 QY 2270 ATGGGTACAGTACACTGCTCAGTGTGATGGTGACCAAAATCTGAGAAATTAACACTA 2329
 Db 220887 TACGGCAGAGTGTATATCTGCTTGGGTGATGGGTGACCAAAATCTGCAAAATCACCAC 220946
 QY 2330 AAGAACTTATCCATGAGAGCAACCACTGTTCCCAAAATCCCAATGAAATTAATAAT 2389
 Db 220947 AAGAACTTATCCATGAGAGCAACCACTGTTCCCAAAATCTGCAAAATTAATAAT 221006
 QY 2390 AATAATTAATAA 2401
 Db 221007 AAAAAACAGCA 221018

RESULT 13

US-09-880-107-3428/C
 : Sequence 3428, Application US/09880107
 : Patent No. US20020142981A1
 : GENERAL INFORMATION:
 : APPLICANT: Horne, Darci T.
 : APPLICANT: Vockley, Joseph G.
 : APPLICANT: Scherf, Uwe
 : APPLICANT: Gene Logic, Inc.
 : TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
 : FILE REFERENCE: 44921-5028-WO
 : CURRENT APPLICATION NUMBER: US/09/880,107
 : PRIOR FILING DATE: 2001-06-14
 : PRIOR APPLICATION NUMBER: US 60/211,379
 : PRIOR FILING DATE: 2000-06-14
 : PRIOR APPLICATION NUMBER: US 60/237,054
 : NUMBER OF SEQ ID NOS: 3950
 : SOFTWARE: Patent In Ver. 2.1
 : SEQ ID NO 3428
 : LENGTH: 99014
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : OTHER INFORMATION: Genbank Accession No. US20020142981A1 U78027
 US-09-880-107-3428

Query Match 15.8%; Score 1358; DB 9; Length 99014;
 Best Local Similarity 78.8%; Freq. No. 1.5e-309;
 Matches 1895; Conservative 0; Mismatches 345; Indels 170; Gaps 16;

QY 131 AGCATATCAAGAAATATATCCACTTGTCAAGTGGTTTCATATACAGGGGTGACAGA-T 189
 Db 88976 AGAAAAAATTAAGATATCCACATGCTCAAGTTGTTTCATATCCAGGATGACAGAT 88917
 QY 190 AGTTTACATACCAAGTCAATTAATGTGTATCATCATCAATTAACAGAAATTAACAAA 249
 Db 88916 GGTTTAATCATGCAAGTTAATCATGTGTATACACCAATTAACAGAAATTAACAAA 88857
 QY 250 ATCAATGATCATCTCAATTAATGATGCTGAAAAAGCAATTTGACAAATTAATTAATTT 309
 Db 88856 ATCAATGATCATCTCAATTAATGATGCTGAAAAAGCAATTTGACAAATTAATTAATTT 88797
 QY 310 TGATTAATTAATCTTACGAAATGACATATGAAGGACATATCTTAATGTAATTAAGCCA 369

Db 88796 TGATTAATTAATCTTACGAAATGACATATGAAGGACATATCTTAATTAATTAAGCCA 88737
 QY 370 TATATGACGAGCCACACGAAACATTAATATGATATGGGAAAAATTGAATTTGCTCC 429
 Db 88736 TCTATGACGAAACCCACAGCCCAATTAATATGATATGGGAAAAATTGAATTTGCTCC 88677
 QY 430 TGAATCTGAAACAAAGCAAGATG-CTACTTTCACCTTCTATTAACATATGATGAG 488
 Db 88676 TGAATCTGAAACAAAGCAAGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 88617
 QY 489 AAGTTTACCCAGAGCAATCAGCAAGAAAGAAATCAAGGAGCCACCAATTAATTAAG 548
 Db 88616 AAGTTTACCCAGAGCAATCAGCAAGAAAGAAATCAAGGAGCCACCAATTAATTAAG 88558
 QY 549 AGGAATCAACTGTCTCTGTTCACTGATGATATGATATGATATGATATGATATGAT 608
 Db 88557 AGGAATCAACTGTCTCTGTTCACTGATGATATGATATGATATGATATGATATGAT 88498
 QY 609 ACTCATCCAGAAAGCTCTGTTCACTGATGATATGATATGATATGATATGATATGAT 668
 Db 88497 ACTCATCCAGAAAGCTCTGTTCACTGATGATATGATATGATATGATATGATATGAT 88438
 QY 669 TAAATGTACAAATCAGTACGCTGTATACCAAGAGTACCAAGTACCAAGTACCAAG 728
 Db 88437 TTAATGTACAAATCAGTACGCTGTATATACCAAGAGTACCAAGTACCAAGTACCA 88378
 QY 729 TCAAGACTCAAACTTTTCAATATGCTGTAA-----AATACTTAAGATATTTCT 782
 Db 88377 TCAAGACTCAAACTTTTCAATATGCTGTAA-----AATACTTAAGATATTTCT 88318
 QY 783 TACCCAGAGAGTGAAGGACCTCTCAAGGAAATCTACAAACACAGCT---GACATCA 838
 Db 88317 TACCCAGAGAGTGAAGGACCTCTCAAGGAAATCTACAAACACAGCT---GACATCA 88258
 QY 839 TAGATGACAAACAAAGTGAAGCAATCCATGCTCATGATGATGATGATGATGATGAT 898
 Db 88257 TAGATGACAAACAAAGTGAAGCAATCCATGCTCATGATGATGATGATGATGATGAT 88198
 QY 899 TGAATGACAAATGTCGAAAGCAATCTACAGTTCATATGATGATGATGATGATGAT 958
 Db 88197 TGAATGACAAATGTCGAAAGCAATCTACAGTTCATATGATGATGATGATGATGAT 88138
 QY 959 CATCATCATCTTCAAGAGCAATGAAAGCAATGAAAGCAATGAAAGCAATGAAAGCAAT 1018
 Db 88137 CATCATCATCTTCAAGAGCAATGAAAGCAATGAAAGCAATGAAAGCAATGAAAGCA 88085
 QY 1019 CAAAAAAG 1078
 Db 88084 CAAAAAAG 88033
 QY 1079 GAGGATCATCATTTACCATCTTCAATCTATCTACAGGCTATATCATCAAAACATCAT 1138
 Db 88032 GAGGATCATCATTTACCATCTTCAATCTATCTATCTATCTATCTATCTATCTAT 87973
 QY 1139 GGCATGACATTAATTAAGCAATGATGATGATGATGATGATGATGATGATGATGAT 1198
 Db 87972 GGCATGACATTAATTAAGCAATGATGATGATGATGATGATGATGATGATGATGAT 87913
 QY 1199 AGCCAAATTAATTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 1258
 Db 87912 AGCCAAATTAATTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 87853
 QY 1259 AAGACATTTCTAGTTAAACAAATGCTGATGATGATGATGATGATGATGATGATGAT 1318
 Db 87852 AAGACATTTCTAGTTAAACAAATGCTGATGATGATGATGATGATGATGATGATGAT 87793
 QY 1319 AACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1378
 Db 87792 AACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 87734
 QY 1379 AACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1438


```

Db 87733 AATCTAGACCTGAACCTATTAAATTTCTAGAGATAACATATAGAAAAACCTTCTAGAC 87674
Qy 1439 ATTCACTTAGGCAAAAGCTTCATGCGCAAGAACCCAAAAGTAAATGCAACAAAAA 1498
Db 87673 ATTGGCTTAGGCAAGAACTTCATGACCAAGAACCCAAATGCAATGCAATTAACAAAG 87614
Qy 1499 ATAAATGATAGAGCTTAATTAACCTAAAAAGCTTTGGCGAGCAAAACATCATTAAGC 1558
Db 87613 ATAAATGCTGGGAC---TTAAACTAAAGCTTTTGGCAAGAAAAGAAAGAGTCAAGC 87558
Qy 1559 AGAGCAACAGACACACCAGCGAGTGAAGAAATCTTCAACAACTAGACATCTGACTAA 1618
Db 87557 AGTGTAA--AGAGAACCCACAGAGTGGAGAAAAGCTTCAACATCTATCATCTGACAA 87500
Qy 1619 GGACTAATATCCGGAATCCACAGAACTCAACAAATCAGCAAGAAAGAAAGCAACAT 1678
Db 87499 GGAATATATCTCAAGATCTCAATGAAGTCAACAAATCAGCAAGATTAACAAACAT 87440
Qy 1679 CCCATGAAGAGTGGCTTAAGAGACATGAATGACATTTCTCAAAAGAAAGATTAACATG 1738
Db 87439 TCCATCAAAAGAGTGGCTTAAGAGACATGAAGAAATCTCAAAAGAAAGATTAACATG 87380
Qy 1739 GCCAAACAGAGAAATATCTTAACATCACTAATGATTAAGGAATGTAAATCAACAC 1798
Db 87379 -----CATATGAAAAATCTCAACATCAACATGATGCGGAAATGCAATCAAAAC 87327
Qy 1799 TGTATGCGATACCACTTACTCTCGAGAGATGTCATTAATTTAAATCTAAATTA 1858
Db 87326 TGCATGCAATACATACCTTACTCTCGAGAGATGCGCAATATCAAAATCAGAAAACAG 87267
Qy 1859 TAGATGTTGGTGGTCTGTGTGTAAAGAACTTTTACATCTCTGTGTGGAAATGTA 1918
Db 87266 TAGATGTTGGTGGTGTGTGTAAAGAACTTTTAA--TGTGCTGTGGGAATGTA 87208
Qy 1919 ACTTGCSCAACCACTATGAGAAAAGTGTGGAATTTCTTAAGAACTTAAGATAGATG 1978
Db 87207 ACAAGTATGGCCACCATGGAAGAAAGTGTGAGATTTCTTGAAGACTGAAGATGAAGT 87148
Qy 1979 ACCATTTGATTCAGCAATCCCATTA----- 2003
Db 87147 ACCATTTGATTCAGCAATCCCATTAAGTGTATCTACCGGAGAAAAGATGCTTATA 87088
Qy 2004 ----- 2003
Db 87087 CGAAGAAATATTGCAACACGTTTATAGAGCAAAATTCGAATTCGAATTCGTGA 87028
Qy 2004 -----AATATGTAATTAATATATATATTTATTTATATCA 2034
Db 87027 ACCAACCCAAATGCCCATCAATCAACAGTGGATTAATTAACGCTGTATATCTATCA 86968
Qy 2035 TGGAAATCAACTCAGCCATTAATAAAGATTAATGATGACATTCACAGCATCTAGATG 2094
Db 86967 TGGAAATCTCCACCCATTAATAAAGATTAATGATGATTAATGATTCACAGCCATCTG 86908
Qy 2095 AATTGGAAGCCCTTATTTCTAAGTGGGTAACTTCAGAAATGAAAAACCAACATCATAT 2154
Db 86907 GATTGGAAGCAATTAATTAAGTGAAGTAACTCAGAAATGAAAAACCAACCTTGATATG 86848
Qy 2155 TCTCACTTAACAAGTGGGGGCTAAGCTGTGAGGACACGAGGCAATTAAGATTAATG 2213
Db 86847 TCTCACTCAATAGTGGAGCTTAAGCTTAAGAGATGCAAGGCAATTAAGATTAATG 86788
Qy 2214 ACTCTGGGACT--TGAAGGGAAGATGGAAGAGAGGAGGATTAAGATTAAGATTAAG 2272
Db 86787 ACTTTGAGACTCGGGGGAAGGAGTGGGAAGAGAGGATTAAGATTAAGATTAAG 86728
Qy 2273 GGTACAGTGTACATCTCTCAGGTGATGAGTCAACCAAAATCTCGAAAT--TACCATTA 2331
Db 86727 GGTTCAGTGTATATCTGCGGGGTATGGGTGCAACCAAAATCTCGAAATGAAACATTA 86668
Qy 2332 GAACTTATCATGAGCAAAACCACTGTTTCCCAAAATCCCAATGAATTAATAATTA 2391
Db 86667 GAACTTATCATGATGAACCAATACCACTGTTTCCCAATGAATTAAGAAAAAACA 86608

```

```

Qy 2392 TAATATTA 2401
Db 86607 AAAACAAACA 86598

RESULT 14
US-10-775-169-149
; Sequence 149, Application US/10775169
; Publication No. US20040175743A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Butczynski, Michael
; APPLICANT: Twine, Natalie
; APPLICANT: Dörner, Andrew
; APPLICANT: Trepicchio, William
; TITLE OF INVENTION: Method for Monitoring Drug Activities In Vivo
; FILE REFERENCE: AM101080 (031896-013000)
; CURRENT APPLICATION NUMBER: US/10/775.169
; NUMBER OF SEQ ID NOS: 5278
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 149
; LENGTH: 136284
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-775-169-149

Query Match 15.6%; Score 1346; DB 17; Length 136284;
Best Local Similarity 76.6%; Pred. No. 1.3e-106;
Matches 1985; Conservative 0; Mismatches 390; Indels 215; Gaps 19;

Qy 1 TGTATGAGCCAAATGTCATCTTAATACCAAAACCGAGAAAGATATA--CAAAAAGAA 58
Db 83890 TGTATGAGCCAAATGTCATCTTAATACCAAAACCGAGAAAGATATA--CAAAAAGAA 83949
Qy 59 ACTTAAACCAATGTCATCTTAATACCAAAACCGAGAAAGATATA--CAAAAAGAA 118
Db 83950 ACTTAAACCAATGTCATCTTAATACCAAAACCGAGAAAGATATA--CAAAAAGAA 84009
Qy 119 ACCCAATCCAAACGATATCAAGATTAATCAACATTTGCAAGTGGTTTCAACAG 178
Db 84010 ACTCAATCCAAACGATATCAAGATTAATCAACATTTGCAAGTGGTTTCAACAG 84069
Qy 179 GGGTGCAGG--ATAGTTAATACATACAGTCAATTAATGTATATCATCATTAACAGAA 237
Db 84070 GGAATGCAAGATGGTTTAAATACAGATGAATTAACGATATCAACATTAACAGAA 84129
Qy 238 TTAATAAC--AAATATCATGATCATCTCAATGATCTGAATAAAGCAATTTGACAAATC 296
Db 84130 TTAATAAC--AAATATCATGATCATCTCAATGATCTGAATAAAGCAATTTGACAAATC 84189
Qy 297 TAACATTTCTTATGATTAATAACCTTCAGCAAAATGCAATTAAGAGCAATTCCT-- 353
Db 84190 CAGATCCCTTATGATTAATAACCTTCAGCAAAATGCAATTAAGAGCAATTCCT-- 84249
Qy 354 -----AATGTAATAAAGCATATATGACGAGCCACAGCAAAATTAATCTGATG 405
Db 84250 TGTATTAATAATGTAATAAAGCATATATGACGAGCCACAGCAAAATTAATCTGATG 84309
Qy 406 GGGAAAGTTGAAACATTTGCTC-----CTGGAATCGAATC 442
Db 84310 GGGAAAGTTGAAACATTTGCTC-----CTGGAATCGAATC 84369
Qy 443 AAGACAGATG--CTACTTCAACATCTTAATCAATGATGAGTGGAGTTTACGAG 501
Db 84370 AAGACAGATG--CTACTTCAACATCTTAATCAATGATGAGTGGAGTTTACGAG 84429
Qy 502 AGCAATCAGCAAGAGAAAGAAATCAAGGCAACCAATCAATTAAGAGAGTCAAACT 561
Db 84430 AGCAATCAGCAAGAGAAAGAAATCAAGGCAACCAATCAATTAAGAGAGTCAAACT 84489
Qy 562 GTCCCTGTCACTGATATATGATTTGATTAATCTAAGAAAACCTTAAGATCATCAGAAA 621

```


Query Match	15.2%	Score 1311.6	DB 9	Length 1503841
Best Local Similarity	76.3%	Pred. No. 8.4e-298		
Matches 1937, Conservative	3	Mismatches 432	Indels 168	Gaps 20

OY	359	AAATTAAGCCATATATGACGAGCCCAACACAAACATTAATACGAATGGGGAAAGTTGAA	418
Db	546349	AATTAATAACCATCTATGAGAGCCCAACACCAACTTATATCTGAAGAGGAAATGGGAA	546408
OY	419	AACATTTGCCCTGAGAACTGGAAACAAGACAGAGATGCTACTTTCACACTTCTATTCAC	478
Db	546409	AGCATTTCCACTTAAGAACTGGAAACAAG-----GATGCCACTTTCACTTCTATTCAC	546466
OY	479	ATATGTAATGGAAGTTTATGACCCAGAGCAATCAGACAAAGAAAGAAATCAAGGCCACCCCA	538
Db	546465	ATCTGATACAGGAATCTTACCCAGAGCAATCAAGACAAAGAAATTTAAGACATCCAA	546524
OY	539	ATCAATTAAGAAGGAAGTCAAACTGTGCCCTGTTCACTGATGATATGATTTGATACCTAGAA	598
Db	546525	AATGTGTAAGAAGACATCAAAATTTGTCGTGTTTGAGATGATATGATGTTGTTACTGAG-A	546583
OY	599	AACTTAAGAACTCATCCGAAAGCTCTTGAACTGTATACATTAATTCAGTAAGATTTCA	658
Db	546584	AACTTAAGATTTCCACCAAAAGCTCTTGAACTGTATTAAGTAATTCAGCAAGTTTCA	546643
OY	659	GGATTCAAACTTAAGTACACAAATCAGTAGACCTGTATACCAACAGTGCCAAAGCT	718
Db	546644	GGATTCAAAAATTAAGTAA-AAAATTAGAGCTCTGTATACCTAACACAGCAACCAAGCT	546702
OY	719	GAGATTCAAATCAAGAACTCAAACTTTTACATATAGCTGTAAA-----AAATACTTAA	773
Db	546703	GAGATTCAAATCAAGAACTCAAACTTTTCACAATATCAGCAAAATAATTAATACRTAG	546762
OY	774	GAATTTCTTACCCAAAGAGGTGGAAGAACTCTACAGGAAAATCTCAAAACACAGCT--	831
Db	546763	GAATTTACTTAACAAAGAGTGGAAAGAACTCTTACAGGAAAATCTCAAAACACCTGTGA	546822
OY	832	--GACATCATATAGATACACAACAAGTGGAAACAATCCCATGCTCATGAGATGGGTAGAA	889
Db	546823	AAGAAATCATAGACACACAAACAAATGGAACAATCCCATGCTCATGAGATGGGT-GAA	546883
OY	890	TCAATATTGTGAAAAAGACATATTGTCCAAAGCAATCTACAGTTCAATGCAATTTCCA	949
Db	546882	TTAAATTGTGAAAAAGACACATAGTCCAAAGCAATCTACATTTCAATGCAATTTCCCT	546941
OY	950	CCAAATATCATCATCTCTTCAACAGACTTGAAAAAACAATCTTAATTCATATG	1009
Db	546942	TCAAATATCATCATCTCTTCAACATTAATGAG-ACAAAACAATCTTAATTCATATG	547000
OY	1010	AACAACAACCAAAAAAAAAAAAAAAAAAACCCTGATGCCCCAAACAGACTTAGCAAAAGCA	1069
Db	547001	AAC-----AAAAAAAAAGCCTGTATGACAAACCAAGACTTAAGC-CAAAAG	547045
OY	1070	ACAAATCTGAGAGCATCATTTACCCATTTCAAACTATATCTACAGAGCTATATATCA	1129
Db	547046	ACAAATCTGAGAGCATCATTTATCATATGTAACAACTAATTTATTAAGGCAATATGACCA	547105
OY	1130	AAACATCATGAGCATGACATTAACCTAAGCAATGACATGACCAATGAAAAAGAGAGAAATC	1189
Db	547106	AAACGCACTGATCTGATCTGATTAATAATGATGATGACCAATGGAACAGATATGAGAAC	547165
OY	1190	CAGAAATAAAGCCAAATATATTATGACCACTGATTTTGGACAAGAGCAAACAAAACATTA	1249
Db	547166	CAGAAATTAACCCAAATACTTTCACCCCACTGATATTCACCAAGAGCAAACAAAAC-AA	547222
OY	1250	AGTGGGGAAGAAAGACATTTCTAGTTAACAAATGATGCTGAGATTAATTTGGCAAGCCACATG	1309
Db	547224	AGTGGGGAAGAGACCCCTATTCACAAATGATGCTGAGATCAATTTGGCAAGCCACATGTA	547283
OY	1310	GAAAGATGAACCTGATCCCTTGTCTCTCACTTAATATCAAAATTTGATACAGAGTGAATC	1369
Db	547284	GAGGAATGAACCTGAT--CTTCATCTCTTATCTATGCAAAATATCAACCTCAAGATGATC	547342
OY	1370	AAAGACTTAATCTGAGACTTAACCATTAATAATTTCTGAAGATTAACATCGAAAAATG	1429
Db	547343	ATATCTTAATCTTAAGATGTGAAACATTAAGATTTTGAAGATTAACATCGAAAAAAC	547402

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 2, 2004, 00:16:09 ; Search time 22711 Seconds
(without alignments)
13833.973 Million cell updates/sec

Title: US-09-867-570-3
Perfect score: 8622
Sequence: 1 tgcataagacatgtcact.....gtccccaagccctacc 8622

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 1821985908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_est4: *
5: gb_est5: *
6: gb_est6: *
7: gb_est7: *
8: gb_est8: *
9: gb_est9: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020.8	11.8	2703	8 AF101616	AF101616 AF101616
2	877.6	10.2	2735	3 HSM804825	AL833512 Homo sapi
3	814.6	9.4	5133	3 AY539945	AY539945 Rattus no
4	801.8	9.3	3654	3 AY387074	AY387074 Rattus no
5	801.8	9.3	4071	3 AY539883	AY539883 Rattus no
6	790.6	9.2	4356	3 AY539886	AY539886 Rattus no
7	770	8.9	2488	3 BC038425	BC038425 Homo sapi
8	769.2	8.9	2231	3 HSM804900	AL833512 Homo sapi
9	749.2	8.7	6966	3 AY539894	AY539894 Rattus no
10	747.8	8.7	3429	3 AY539894	AY539894 Rattus no
11	744.8	8.6	8575	3 CR627265	CR627265 Homo sapi
12	720.4	8.4	3232	3 AK038045	AK038045 Mus muscu
13	708.2	8.2	889	6 CA455045	CA455045 AGENCOURT
14	702.8	8.2	2776	3 BC028293	BC028293 Homo sapi
15	683.4	7.9	871	7 CN839125	CN839125 AGENCOURT
16	672.2	7.8	3315	3 AY387051	AY387051 Rattus no
17	663.2	7.7	863	7 CN835944	CN835944 AGENCOURT
18	655.8	7.6	2354	3 AK052675	AK052675 Mus muscu
19	645	7.5	5064	3 AY539942	AY539942 Rattus no
20	631.6	7.3	3309	3 AY387094	AY387094 Rattus no
21	627.2	7.3	1821	8 AF101873	AF101873 AF101873
22	627	7.3	1897	3 AY325175	AY325175 Rattus no
23	620.2	7.2	3523	3 BC035112	BC035112 Homo sapi
24	610.6	7.1	3222	3 AY539902	AY539902 Rattus no

25	607.4	7.0	851	7 CN839827	CN839827 AGENCOURT
26	606	7.0	835	4 CN843633	CN843633 AGENCOURT
27	604.4	7.0	764	4 BG198766	BG198766 RST18035
28	603.2	7.0	827	7 CN831885	CN831885 AGENCOURT
29	601.8	7.0	3662	3 HSM804836	AL833523 Homo sapi
30	598.2	6.9	924	7 CN839634	CN839634 AGENCOURT
31	596.2	6.9	3579	3 AK048255	AK048255 Mus muscu
32	589.4	6.8	5145	3 AY539881	AY539881 Rattus no
33	585.8	6.8	688	2 BE439409	BE439409 HTML-025F
34	585.4	6.8	899	7 CN843534	CN843534 AGENCOURT
35	582	6.8	835	7 CN834097	CN834097 AGENCOURT
36	581.8	6.7	826	7 CN839711	CN839711 AGENCOURT
37	567.4	6.6	2855	3 AK080274	AK080274 Mus muscu
38	560.2	6.5	699	8 AQ286739	AQ286739 RPT11-89
39	557.4	6.5	825	7 CN835770	CN835770 AGENCOURT
40	551.6	6.4	1076	8 B2611142	B2611142 WHAC163TR
41	548.2	6.4	787	4 BG208126	BG208126 RST27616
42	546	6.3	855	7 CN839609	CN839609 AGENCOURT
43	544.6	6.3	825	7 CN843258	CN843258 AGENCOURT
44	544.2	6.3	9860	3 HSM80687	BK640640 Homo sapi
45	540	6.3	4583	3 CR627206	CR627206 Homo sapi

ALIGNMENTS

RESULT 1	AF101616	2703 bp	DNA	linear	GSS 06-NOV-2000
LOCUS	AF101616				
DEFINITION	AF101616 Human Homo sapiens genomic clone pTMB233.328P6, genomic survey sequence.				
ACCESSION	AF101616				
VERSION	AF101616.1	GI:4249719			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Beppler, G., O'Brian, K.C., Kim, Y.C., Schreiber, G. and Pitterle, D.M.				
TITLE	A 1.4-Mb high-resolution physical map and contig of chromosome segment 11p15.5 and genes in the LOH1A metastasis suppressor region				
JOURNAL	Genomics 55 (2), 164-175 (1999)				
MEDLINE	99134294				
PUBMED	9933563				
COMMENT	Contact: Beppler G Medicine and Radiology Duke University Medical Center Box 2610, MSRB, Room 117, Durham, NC 27710, USA part of a 1.4 megabase contig including the LOH1A metastasis suppressor region Bin B Class: unknown.				
FEATURES	Location/Qualifiers				
source	1..2703 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /map="11p15.5" /clone="pTMB233.328P6" /clone_id="Human"				
ORIGIN					
Query Match	11.8%; Score 1020.8; DB 8; Length 2703;				
Best local Similarity	73.6%; Pred. No. 1.8e-216;				
Matches 1486; Conservative 0; Mismatches 497; Indels 37; Gaps 13;					
Qy	1 TGTATGAGCCATGTCCTTATATACCAAGCAAGAGATATA--CAAAAAGAAA 58				
Db	91 TTTATGAGCCACATCATCTGATACCAAGCCTGCAGACAGACAACAAAAAGAGA 150				
Qy	59 ACATATGACAGTACACATGATATATATACATGACGAAATCCCAAAATATCTAGCTA 118				

Db 151 ATGTTACACAAATATCTTGATGAACATGATGCAAAAATCTCATATAAATCTGGCA 210
 Qy 119 ACCCAATCAACAGATATCAAGAGATATCCACCATTTGCAAGTGGTTTCAATCCAG 178
 Db 211 ACTGATTCAGAGACATCAAAAAGCTTATCCACCATGATCAAGTGGCTTATCTCCG 270
 Qy 179 GGGTGC -AGSANTAGTTAACTATCAACAAGTCAATTAATGATGATGATCAATAAGAA 237
 Db 271 CCATGCAAGGCTGGTTCAACATACGAAAATCAATTAATCAATGACATATTAACGAA 330
 Qy 238 TTAATAAATAAATCAATGATCATCTCAATAGATCTGTAATAAAGCATTTGACAAATCT 297
 Db 331 CCAAGAGACAAACCATATGATTTCTCAATAGATCAGAAAAGGCTTTGACAAATTC 390
 Qy 298 AACATTTCTTATGATTAATAACCTTCAGCAAAATGACATAGAAAGACATCTTAATG 357
 Db 391 AACATGCTTCAATGCTTAATAAATCTCAATTAATTAATGATGGAATATCTCAAAA 450
 Qy 358 TAAATTAAGCCATATATGACGACCAAGCAAACTTAATATGGAATGGAAGAAATTTGA 417
 Db 451 TAAATAGAGCTATCTATGACAAACCAAGCCAAATATCAATGATGAGCAAAAATCTGG 510
 Qy 418 AACATTTGCTGCGAATCTGGAACAAGCAAGATG -CTACTTGCACCACTTCTATTTGA 476
 Db 511 AAGCATTCCTTTGAAAACTGGCACAAGGCAAGGATGCCCTCTCTACCACTCTTATTTGA 570
 Qy 477 ACATATGATGGAAGTTTATGACCAAGCAATCAACAAAGAAAGAAATCAAGGCAACC 536
 Db 571 ACATATGATGGAAGTTTATGACCAAGCAATCAAGGCAAGAAATTAAGGCAATTC 630
 Qy 537 AATCAATTAAGGAAGTCAAACTGCTGCTGTCATGATGATGATGATGATGATGATGATG 596
 Db 631 AATTAAGAAAAAGG -AGTGAATATGCTGCTGTTGACATGATGATGATGATGATGATG 689
 Qy 597 AAAACCTTAAGACTCATCCAGAAAGCTCTAGAACTGATATACATAATTCAGTAAAGTT 656
 Db 690 AAAACCTTAAGACTCATCCAGCAAAATCTCTTAAGTGAATGAACCACTTCAAGAAAGTCT 749
 Qy 657 CAGATATCAAACTAAATGTTACAAATCAATGATGATGATGATGATGATGATGATGATG 716
 Db 750 CAGATATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 809
 Qy 717 CTGAATCAATCAATCAAGACTCAAACTTTTACATAGCTGTAAT -AAAAATCT 770
 Db 810 CAGAAAGCCAAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 869
 Qy 771 TAAAGATATTTCTTACCAGAGAGTGAAGGACTCTTACAAAGAAAACTTACAAACACAGC 830
 Db 870 TAGGATTCACACTTACAGAGATGTAAGGACTCTTACAGAGAACTTACAAACACAGC 929
 Qy 831 ---TGACATCATAGATGACAAACAGTGAAGACATCCCATGCTCATGATGATGATGATG 886
 Db 930 TCAATGAATTAATAAGAGATATCAAAATGAAAGAAATTCATGCTTATGATGATGATGATG 989
 Qy 887 GAATCAATTTGTTGAAATGACATATGTCGCAAAAGCAATCTTACAAAGTTCAATGCAATTC 946
 Db 990 GAATC -ATATGCTGAAAAATGCTATGCTGCAAGGATTAATTTATGATGATGATGATGATG 1048
 Qy 947 CCACCAAAATATCATCATCTTCTTACAGAACTTGAAGAAAAAACTTCTTAATATTCAT 1006
 Db 1049 CCATCAAGCTTACCAATGATCTTTCTTACAGAACT -GGAAGAAAACTTCTTAATATTCAT 1107
 Qy 1007 TGGAAACAACCAAAAAAAGAAAAAAGGCAATGATGATGATGATGATGATGATGATGATG 1066
 Db 1108 TGG-----AATCAAAAAAGAGCCACATCAACCAAGCAATCTTAAGCCAAA 1153
 Qy 1067 AGAACAATCTGAGAGGATCACTTACCATCTTCAAACTTATCTTCAAGGCTTATATCA 1126
 Db 1154 AGAACAAGCTGAGAGGATCACTGCTTCACTTCAAACTTATCTTCAAGGCTTATATCA 1213
 Qy 1127 CCAAAACATCATGAGCTGATCATTAATACTGAGCATATGAGCAATGAGAAAGAGAGA 1186
 Db 1214 CCAAAACATCATGAGCTGATCTGATACCAAAACAGAGATCTTACCAATGAGAAAGAGAGC 1273

Qy 1187 ATCCAAAAATAAGCCAAATTAATTTATAGCAATGATTTTGAACAAAGCAAAACA 1246
 Db 1274 CCTCAAAATTAATGCGGATATCTACAACTATCCATCTTTGACAAACCTGAGAAACA 1333
 Qy 1247 TAAAGTGGGAAAGACATTTAGTTAACAATGATGATGATGATGATGATGATGATGATG 1306
 Db 1334 AGCAATGGGAAAGATTCCTTATTTAATATGATGATGATGATGATGATGATGATGATG 1393
 Qy 1307 GTGAAAGAAATGAATCGATATCCCTTGTCTCTCATTTAATACAAATTTGATACAAAGATG 1366
 Db 1394 GTAGAAAGCTGAATCGATTC -TTCTTACACCTTATACAAAAATTAATTTCAAGATG 1451
 Qy 1367 ATCAAAAGCTTAAATCTGAGACCTTAAACATTAATAATTTAGAAATTAATGATGATG 1426
 Db 1452 ATTAAGACTTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1511
 Qy 1427 ATGCTTCTGACATTCATTTAGGCAAGACCTTATGATGATGATGATGATGATGATGATG 1486
 Db 1512 ACCATTCAGAGCATATGAGCATGAGGCAAGGATCTTATTAATAACCAAAAGATGCA 1571
 Qy 1487 ACAAAAAATAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1546
 Db 1572 ACAAAAAATAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1631
 Qy 1547 ACATCATTAAGAGAGCAACAGACACCCAGGATGATGATGATGATGATGATGATGATG 1606
 Db 1632 GAATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1691
 Qy 1607 GCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1666
 Db 1692 TCATCTGACAAAGGCTAATATTCAGATCTTATGATGATGATGATGATGATGATGATG 1750
 Qy 1667 AAAGCAACATCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1726
 Db 1751 AAAGCAACATCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1810
 Qy 1727 GATATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1784
 Db 1811 GACATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1870
 Qy 1785 ATGTAATCAACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1844
 Db 1871 ATGCAATCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1930
 Qy 1845 AAATCTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1903
 Db 1931 AAGTCAGAAACACAGTGTGAGAGATGATGAGAAACAGAAACATTTTACATG 1990
 Qy 1904 CTGATGAGAAATGTAATCTTGGCCACCACTATGAGAAACAGTGTGAAATTTCTTAAGA 1963
 Db 1991 TTGATGAGAAATGTAATCTTGGCCACCACTATGAGAAACAGTGTGAAATTTCTTAAGA 2050
 Qy 1964 ACTAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2003
 Db 2051 TCTAGAACTGAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2090

RESULT 2
 HSN804825 2735 bp mRNA linear HTC 03-AUG-2004
 LOCUS Homo sapiens mRNA, cDNA DKFZp686B0737 (from clone DKFZp686B0737).
 DEFINITION AL833512
 ACCESSION AL833512 GI:21734156
 VERSION AL833512.1
 KEYWORDS Homo sapiens (human)
 SOURCE HTC
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 2735)
 Bahr, A., Lauber, D., Mewes, H.W., Weil, B., Amidi, C., Osanger, A.,
 Fobo, G., Han, M., and Wiemann, S.
 The German cDNA Consortium

TITLE	Direct Submission
JOURNAL	Submitted (13-JUL-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

FEATURES

Location/Qualifiers

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp686E0737"
/db_xref="taxon:9606"
/clone="DKFZp686E0737"
/clisue type="cantis"
/clone_lth="666 (synonym: hlcc3) . Vector pSport1_5fl; host
DH10B; sites SfilA + SfilB"
/dev_stage="adult"
note="unspliced mRNA"

```

ORIGIN

Query Match	10.2%;	Score 877.6;	DB 3;	Length 2735;
Best Local Similarity	69.6%;	Pred. No. 1.7e-184;		
Matches 1401; Conservative	0;	Mismatches 569;	Indels 42;	Gaps 14;

Qy	1	TGATGAGCGCAAGTCACCTTAAATACCAAAACGAGAAAGATATACAAAAAGAAAC	60
Db	524	TCACACGGCCAGATTAATACGTGTGATACCAAAACGAG--ATTAAAGCAGATCAGAGAAAC	581
Qy	61	TATGACCCAGTACCACTGTATGAATATACATGACGAAATCCCCACAAATATCTGTAC	120
Db	582	TACGGGCGAATATATCTGTATGAATAACTTAATCAAAAAATCTTAGCAAAATCTGCAAC	641
Qy	121	CCAATCCAAACGACATATCAAGAAGATATCCACCATTGTCAAGTGGGTTTCAATACCAAG	180
Db	642	CAAAACCAACACACATTTAAABGTCACTATCATGTGCGCAAGTGAATTTATTTCAAGG	701
Qy	181	GTCG-AGGATAGGTTAACTATCAACAAGTCAATTAATGTATACATCACTAAACGAATT	239
Db	702	ATGCAAAAGATGCTTCAACATACAAATCAATCATGTGACACATCATATCAACGAAATG	761
Qy	240	AAAAACAAAATACATGATCATCTCAATATATGCTGAAAAAGATTTGACAAAATCTAA	299
Db	762	AAGACCAAAATTCATATATGTACTTCAATATGATGCTGAAAAAGCATTTGATTAATTTCC	821
Qy	300	CATTTCTTTATGATTAAACCTTGACGCAAAATGCAATAGAAAGACATACCTTAATGTA	359
Db	822	CATCTTTTCATGATAAAAATCTTAAAAATCTGGGTAGAGAAAGCATATCTCAACACA	881
Qy	360	ATTAABCCATATATGACGGGCCACAGCAAACTTATATCTGAATGGGGAAAAAGTTGAA	419
Db	882	ATTAABGTCTATTAABAAACAGCCACAACTATGTATCATATGAAATGGGAAAAAACTGAA	941
Qy	420	ACAATTGCTCTGAGACTGGAACAGACAGACAGAGATGCT-ACCTTACACCACTTTATTCAC	478
Db	942	GACTTCTCTTAATCATCTGTAGAAAGACAAAGACGCTCACTTTACCTGTATTTCAAC	1001
Qy	479	ATATGTAATGAAGTTTATGACGAGCGAATGACACAAAGAAAGAAATTAAGGCAACCA	538
Db	1002	ATTGCACTGAGCGTCTAGCTGAGCAATTAAGACAAAGAAAGAAATTAAGAAATTTCA	1061
Qy	539	ATCATATAAGGAGAGTCAAATCTGCTCCGTTCACCTGATGATGATGATATTAATCTAGAA	598
Db	1062	ATTGAAAGAGAAAGACAAATTAATCTCTGTGTTGTAGATTAATTAATCTTAATTTGGA	1121
Qy	599	AACCTTAAGATCATCGAAAGCTCTAGACTGATACATTAATTCAGTAAAGTTTCA	658

D	1122	AAACCCAAAGCTCCACAGAAAAACTC-----AGCACTGATTAATTCAGTAAGTTGCA	1175
O	659	GGATACAAACTAAATGTACACAATTCAGTAGCACTGCTATACACCAACGTGACCAAGCT	718
D	1176	GGATACAAAATCTCATCTTAATAAATGAGTAGCAGTTCTATATGTCAACACCAAGCAAGCT	1235
O	719	GAGAAATCAATCAAGAACTCAACACTTTTTCATTAAGTGTAAAAAATACTTAAGATA	778
D	1236	GAAGAAAGAAATTAAGAAAGTATATCCATTCACAAATACCTACAAATACTGTAAAAATG	1295
O	779	-TTCTTACCCAGAGGTGAAGACCTCTACAGAGAAACCTACAAAAACACAGCT----GA	833
D	1296	GGAAATTACCAAGAAAGTAGAAGATCTCTACATGAAAAATGTAAAAATATTGATTGAAGA	1355
O	834	CATCATAGATGACACAAACAGATGAAAAACATCCCATGCTCATGATGGGTAGATCA	893
D	1356	AATCAAGAGGACACAAAAAATGGAAGAATATCCATGTTTATAGATTGGAAGATCA	1415
O	894	TATTGTGAAAAATGACCATATTGCCAAAAAGCATCTACAGTTCAATGCCAATCCGATCA	953
D	1416	TATTGTAAAAATATCCACACTACCCAAAGCATCTACAGATTCAATCAATCCGATCA	1475
O	954	AATATCATCATCATCTTCTCACAGAACTGAAAAAAAACAATCTTAAATTCATATGAAACA	1013
D	1476	ACTACCGATACATTTCTTACCCGAAATAG-AAAAATTAATTTTAAATTTATATGAATC	1533
O	1014	ACAACCAAAAAAAAAAAAAAACCCTGATGCCAAAGCAGACCTTAGCAAAAAAGACAA	1073
D	1535	AC-----AAAAGACCAGATAGCCAAAGCTAGCTTGAACAAAAAGAACAA	1580
O	1074	ATCTGAGGCACTCATTTACCATCTTCAAATCTATCTACAGGCTATATACCCAAAC	1133
D	1581	AACGTGTGAAATCGTATTAACCTCTGCACTTCAAATTTACTACAGAGCTGTAGTAAACAAAC	1640
O	1134	ATCATGCACTGACATTAACACTTAGCGCATAGACCAATGAAAGAAGAGAAATTCAGA	1193
D	1641	AGCATGTATTGGCATTAAGATTAACAATAGACCAATGAAACGAATATAGTAACCCAGA	1700
O	1194	AATTAAGCCAAATATTAATAGCCCACTGATTTTGAACAAAGCAACAAAACTATAAGT-	1252
D	1701	AACCAATACATACATCTACATGCACTCATTTTCAACAAAGTTGCCAAGAACATGACTG	1766
O	1253	GGGGAAAGACATTTCTAGTTAACAAATGTGTGTGATTAATTGGCAAGCCACATGTGAA	1312
D	1761	GGGGAAAGACGGGTCTTCAATTAATGTGTGTGGAAAGCCAGATATCCATATATACATA	1820
O	1313	GAATGAACCTGATCCCTTGTCTCTCACTTAATCAAAAAATTGATATACAGATGATCAA	1372
D	1821	GAATGAACCTGATCCCTTGTCTCTCACTTAATCAAAAAATTGATATATATATATATAT	1879
O	1373	GACTTAATATCGAGACCTTAACCACTAATAATTCTAAGAGTAACTACAGAAAAATGCTT	1432
D	1880	GACTTAATATATAGACTGTAACCTGAAATTTATGACACATGGAAGAGACATTTAGAGAA--CTC	1937
O	1433	CTAGACATTCATCTTAGGCAAGACCTTATGCGCAAGAACCCAAAGTAAATGCAACAAA	1492
D	1938	TCTGACTGTGACTGGGCAAGAGTGTCTGTGTATATTCACAAATCATATAGCAACCAAG	1997
O	1493	ACAAAATTAATATAGATAGACTTAATTAATACTAAAAAGCTTTTGGCAGCAAAAAACATC	1552
D	1998	CAACAAATATGCAAAATAGATCTCTTCAAGTTGAAAGCTTGTGCACAGCAAAACA----	2052
O	1553	ATTAGCAGAGCAAAACAGACAAACCCAGAGTAGAGAAAAATCTTCAACAACTAGACATCT	1612
D	2053	ATCAACAAATATGAAAGAGGACCAACCAAGAACTGAAAGAAATATTGGCAACTACTCATCT	2112
O	1613	GACTTAAGACTAATATATCCGAAATCCACAAAGAACTCAACAAATCAGACAAAGAAAGCA	1672
D	2113	AACAAATGATTTGATTAACAGAAATATATATAGAGAGCTGCAACACATATAGAGAAAAATAT	2172
O	1673	AACAAATCCCATGAAGAAGTGGCTTAGAGACATGATATGACAAATTTCTAAAAAGAAATATA	1732
D	2173	AATATATCAATTTAAA-AATAGACAAAAATATATATGATATGAGGTTCCTCAAAACAAAGCAT	2231

QY 771 TAAGATATTTCTTACCCAGAGGTGAAGACTCTTCAAGAGAACTTACAAACACAG- 829
 Db 3536 TCGGTGACTTAAACAGAGTAAAGATCTGTCAATTAAGACTTACAAAGCACTGA 3595
 QY 830 ---CTGACATCATGATGACACAAACAAGTGAAGAACATCCATGCTCATGATGGTA 886
 Db 3596 AGAAGAAATTTGAAGAGACTCTGAGAGAGTGAAGAGATCTCCATGCTCATGATGGCA 3655
 QY 887 GAATCAATATTTGTAAGAAATGACATATTTGCCAAAGAAATCTCAAGTTCATGCAATTC 946
 Db 3656 GGAATTAATATAGTAAATATGAGCCATTTTACAAAGAAATCTACAGATTCATGCAATTC 3715
 QY 947 CCACCAAAATATATCATCTTCTTACAGAACTAGAAAAAAACAATTTCTAAATTCATA 1006
 Db 3716 CCATCAAAATATCAATCAATCTTCAAGAGTTAG-ACAGAACAAATTTGCAAAATTCATC 3774
 QY 1007 TGGAAACAACAACCAAAAAAAGAAAAAAGCCGATAGCCAAAGCAAGACTTGAACA 1066
 Db 3775 TGGAAATAC-----AAAAAACCCAGAGATAGCTTAAGCTTCTTCAACATA 3820
 QY 1067 AGAACAATCTGAGGATCATCAATTAACCATCTTCAACTACTACAGAGCTTAAATCA 1126
 Db 3821 AAGAGACTTCAAGGGGAAATCATCTATCCCTGAACTCAAGCAATATTAACAGCAATATGCA 3880
 QY 1127 CCAAAACATCATGCTGACATTAACCTGAGCACTATGACCAATGAGAAAGAGAGA 1186
 Db 3881 TAAAAATCGGTGTATTTGTGACAGACAGACAGATAGACCAATGAAATTTGAATTAAG 3940
 QY 1187 ATCCAGAAATTAAGCCAAATATATATAGCCAACTGATTTTGAACAAAGCAAAAAACA 1246
 Db 3941 ACCCAAGAAATGAACCCCACTCTATGCTCACTTGAATTTTGAACAAAGAGCCAAACCA 4000
 QY 1247 TAAAGTGGGAAAAAGACATTTCTAGTTAAACAATGCTGAGATTAATTTGCGACCAAT 1306
 Db 4001 TCAATGGAAGAAAGATAGATTTTTCAGCAATGCTGCTGCTCACTGAGAGTCAACAT 4060
 QY 1307 GTGAGAAATGAAGAACTGATCCCTGCTGCTCACTTAATACAAATTTGATCAAGATGG 1366
 Db 4061 GTAGAAATATGCAAGATGATCATGCT-TATCACCTGTACAAAGCTTAAGTCCAAAGTGG 4119
 QY 1367 ATCAAAAGCTTAATCTGAGACCTAAACCAATAAAATTTCTAGAGATTAACATCAGAAA 1426
 Db 4120 ATCAAGAGCTCCATCACTCAACAGACACTCAAACTAATAGAGAAAGCTAGGAGAG 4179
 QY 1427 ATGCTTGAACATTCCTTAGGCAAGACTTCAAGCAAGAAACCAAAAGTAAATGCA 1486
 Db 4180 CATCTGGAACACATGGGCACTGAGAAAAATTTCTTAACAAACACCAATGGCTTAGCT 4239
 QY 1487 ACAAAAACAAAATTAATAGATGAGCTTAATTAACCTAAAGCTTTTGGCGACGAAA 1546
 Db 4240 CTAAGATCAAGAAATCGAACAATGGAGATCTCAATAAAGCTTCTGTAAGGAGAG 4299
 QY 1547 ACAATCATTTGACAGAGAAACAGCAACCAACGAGTGAAGAAATCTTCAACAACTGA 1606
 Db 4300 GACACTGTGTTGGAACAAACGCAACCAACATGAGGAAAGATCTTTTCACAATCTT 4359
 QY 1607 GCATCTGACTAAGGACTAATATCCGAGATCCACAAAGAACTCAACAATCAGCAAGAG 1666
 Db 4360 ACAAGATAGAGCCCTTATATCCAAATATACAAAGAACTCAAGAAAGTTCAGCCGAG 4419
 QY 1667 AAGCAACAATCCCATGAAGAGAGTGGCTTAAGGACATGAATGACAATTTCAAAAGAA 1726
 Db 4420 GAAACAATTAACCTATTTAAAGAAATGGGGTTCAAGCTTAACAAAGAAATTCACAGCTGAG 4479
 QY 1727 GATATCAATATGCGCAACAAACA--GGAAAAATGCTTAATCATCACTAATGATTAAGGAA 1784
 Db 4480 GAATGCCGAATGGTTGAGAAACACTTAAGAAATGTTCAATCTTTAGCTATAGAGGAA 4539
 QY 1785 ATGTAATCAACACTGTAATGAGATACCACTTCTCTGCAAGAAAGTGTATATTTAA 1844
 Db 4540 ATGGAATCAACAAACCCGAGATTTCACTCAACAGTGAAGATGGCTTAAGATCAAA 4599

QY 1845 AAATCTAAATATATAGATGTTGGTGGGTCTGTGATGATTAAGAGACACTTTTCACTGC 1904
 Db 4600 AACTCAGGGGACAAACATGCTGCGAGATGTGAGAGAAAGAGAAACACTTCTCATTTGT 4659
 QY 1905 TGTGTGAAATGTAACCTTGGCGAACCACTATGGAAGAAACAGTGTGAAATTTCTTAAGGA 1964
 Db 4660 TGTGTGAAATGTAACCTGTAACCACTATGGAAGAAACAGTGTGAAATTTCTTAAGGA 4719
 QY 1965 CTAAAGTATGATGACCACTTTGATCCAGCAATCCC 1999
 Db 4720 TTGACATTAAGTAACTGCTGAGAGATCCAGCTAATACC 4754
 RESULT 4
 AY387074
 LOCUS
 DEFINITION
 Rattus norvegicus LRRGT00088 mRNA, complete cds.
 ACCESSION
 AY387074
 VERSION
 AY387074.1 GI:37361861
 KEYWORDS
 HTC.
 SOURCE
 Rattus norvegicus (Norway rat)
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 3654)
 Xu, C.S., Chang, C.F., Han, H.P., Wang, G.P., Chai, L.Q., Yuan, J.Y.,
 Yang, K.J., Zhao, L.F., Ma, H., Wang, L., Wang, S.F., Xing, X.K.,
 Shen, G.M., Shi, J.B., Rahman, S., Wang, Q.N. and Zhang, J.B.
 Liver regeneration after PH
 Unpublished
 2 (bases 1 to 3654)
 Xu, C.S., Chang, C.F., Han, H.P., Wang, G.P., Chai, L.Q., Yuan, J.Y.,
 Yang, K.J., Zhao, L.F., Ma, H., Wang, L., Wang, S.F., Xing, X.K.,
 Shen, G.M., Shi, J.B., Rahman, S., Wang, Q.N. and Zhang, J.B.
 Direct Submission
 Submitted (10-SEP-2003) Henan Bioengineering Key Lab, Henan Normal
 University, No.148 Jianshe Road, Xinxiang City, Henan 453002,
 P.R.China
 FEATURES
 source
 location/Qualifiers
 1..3654
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 1..3654
 /note="related to liver regeneration"
 /codon_start=1
 /product="LRRGT00088"
 /protein_id="AA091044.1"
 /db_xref="GI:37361862"
 /translation="MDLTDYRTFYPTAKGYTFPSAPHTFSKIDHIIIGQTKGNRYR
 KIHILPCVLSDHGLKLVFNKKGBMTYVLAALNNLNLNLYVEBIIKKEIAXNLEFN
 ENBDTYSNLMQMKAVLRGLIALASCRKORAVYSSITAHILKALBOEAXNPRS
 RROEITLRAEINQVETKRTIERINRKSFPFEKINKIDPLARLTGHEQVQINKI
 RNEKGDITDSEBIIQKIRSYKLVSTKEINQEMDNIFDRYQVSLNEDQINQNLN
 PITKEIEAVIKGLPTKSPDPGFSFAPOTPEIDILPILSKLPHKIEIDGALNSF
 YEATITLIPKHDOTTKENRPLISNIDAKILKILARIQIHITIIIHDOVGLI
 PMOGWENIRKTIIVIHINKLEONNMIIISLAEKXAPLOHPEMLKVLERIIOGP
 YLNIIVAYIKSPANIKLNEBELEIPLKSTGCGELSPYLVYEVYVLAARLROOK
 EINGIOQKEBVKISLPADMTIVLSPKSTIRBLQVNDLYNKNKRTLKELEEDLRW
 LYTEKQAEKEIREETPIIDPNNIKILGVTLTKQVNDLYNKNKRTLKELEEDLRW
 KDLPCSWIGRIINWKAALIPKAIYRFAIPDKIPIOPKEIDRLTCRFINNNKPRIL
 KALINNRKTGGITLIPELKOYRALIVKTAMWYRDIQDMNRIEDPENAPHYGHL
 IPDKAKTIQMKKDSIFSKCMFNMRATGRMOIDPELSCTKLKSKIKIDHLIKPDT
 LKILBEKIGHLIHMGTGKMLKTPAAVALRERIDWDILIKLQSPKADIVVTRR
 OPTDMKIFNPPTDRLGISIKYELKELKDRERENNDIKKWSBLAKEFTABEGRMAE
 KHLKKSTSLVIREMOIKTLRFLHLEVRALAKINSQSDSCWRGSGRTLLHGMDC
 RLKVPFWKSVYRPLRKIDILPEDPALPLGIYPKDASYKRDCTMPLAALFIAR
 KMKEPRCPSTEEYIQKMYIYTWBYSYAIKNBSMKVGMLELENTIISLEIQTISQD
 IHGQVYKTLSTGDIQTLMDCKPSGSGWRGAPLTVLDNTEPFCHEINNDIYAL
 KDIQVIOKEICSELEAAAYALGTBETSHQFQVGSATYDGERPLEHPMKGLIYLVLLTT
 LTAINSQPSPAVY"
 CDS

ORIGIN

Query Match 9 3%; Score 801.8; DB 3; Length 3654;
 Best Local Similarity 65.6%; Pred. No. 1.5e-167;
 Matches 1321; Conservative 0; Mismatches 662; Indels 32; Gaps 9;

1 TGTATGAAGCCCATGCTCACTTTAATACCAAAACGAGAAAGATATAT--CAAAAAGAAA 58
 Db TCTATGAAGCCCATGCTCACTTTAATACCAAAACGAGAAAGATATATATTAAGAAA 1060
 QY 59 ACTATGACACGATACCATGATGATATATATGACAGAAATCCCAAAATATCTAGCTA 118
 Db 1061 ACTTACAGACCAATTTCCCTTATGAAATATGACGCAAAAATATCTCAATTAATTTCTGCA 1120
 QY 119 ACCCAATCCACAGATATGAGAAAGTAAATCCATGTCAGAGGCTTCATATCCAG 178
 Db 1121 ACCGAATTCAGAGACATCAAAACCAATATCCATGATATCAAGTAGGCTTCATCCAG 1180
 QY 179 GAGTGA--GATAGGTTAATATACACAAAGTCAATTAATGATGATACATCAATTAACGAA 237
 Db 1181 GCATGACAGGATGCTTATATATACGAAACCAATCAAGTATCCATTAATTAACAAAC 1240
 QY 238 TTAATAACAAAATCAATGATCATCTCAATGATGCTGAAAAAGCATTTGACAAAATCT 297
 Db 1241 TGAAGAACAAGAACACATGATCATTTCAATGATGCTGAAAAAGCATTTGACAAAATTC 1300
 QY 298 AACATTTCTTATGATTAATACCTTCAGCAAAATGACATAGAAAGACATTAATG 357
 Db 1301 AACACCCCTTCATGATTAATAGTCTGAAAAAGTATGAAATTCAGGCCATTAATCA 1360
 QY 358 TAAATTAAGCCATATATGACGAGCCACAGCAAAACATTAATGAAATGAGGAAAAATGGA 417
 Db 1361 TAGTAAAGCCATATATACAGCAACAGTTGCTAAACATTAATTAATGAGAAACCTTG 1420
 QY 418 AAAACATTTGCTGAGAACTGGAACAAAGCAGAGAG--CTACTTTTACCACTTTATTTCA 476
 Db 1421 AAGCAATCCCATTAATATGAGGACATGACAAAGGCTGCCCATCTCTCTACTTATTTCA 1480
 QY 477 ACATGATGAGTAAAGTTTAAAGCCAGACATGACAGAAAGAAATCAAGGGACCC 536
 Db 1481 ATATGTTCTTAAAGTTCTAGCCAGAGCATGACACAAAGAAATCAAGGGATAC 1540
 QY 537 AAATCAATTAAGAGAGAGTCAAACTGCTGCTGTTCACTGATGATGATGATTAATCTAG 596
 Db 1541 AGATGGAAGAAAGAGAGTCAAAATATCACTATTTGCAATGACATGATGATTAATTTA 1600
 QY 597 AAAACCTTAAGACATCTCAGAAAGCTCTAGAACTGATATCAATTAATTCATGAAAGTT 656
 Db 1601 GTGATCCCAAAAGTTCCACAGAGAACTTAAGGCTGATGAACAACTTCAGCAAAAGTGG 1660
 QY 657 CAGATTAACAACTTAATGTAACAATATCAATGCTAGCACTGCTATACCAACAGTGAACAG 716
 Db 1661 CTGGGATTAATTAATCTCAATTAATATCACTGCTCTCTATACAAAGAGAAACAG 1720
 QY 717 CTGGAATCAAAATCAAGAACTCAAACTTTTACAATAGCTGTAA--AAAAATCT 770
 Db 1721 CCGAAGAAAGAAATTAAGGAAAGCAACCTTCAATATAGACCAAAATTAATTAAGTACC 1780
 QY 771 TAAAGATATTTTAACTCCAGAGAGTGAAGAGCTCTACAAAGAAACTTACAAAACAG-- 829
 Db 1781 TCGGTGTGCTTTAACTCAAGAGTGAAGTGAATCTGTATACATTAAGAACTTCACAGCTGA 1840
 QY 830 ---CTGACATCATAGATGACAAACAGAGTGAAGAAACATCCGATGCTAGTGGGTA 886
 Db 1841 GGAAGAAATTAAGAGAGCTCTCAAGAGATGAGAAAGTCTCCATGCTCAAGATTTGGCA 1900
 QY 887 GAATCAATATTTGAAATGAGACCATATTTGCAAAAGCAATCTCAAGTTCAATGCAATTC 946
 Db 1901 GGATTTAATATGTAAGAAATGGGCATTTTACAAAGAACTTCAAGTTCAATGCAATTC 1960
 QY 947 CCAACCAAAATATCATCATCTTTTCAAGAGAACTAGAAAAAAACAATTTCAAAATTCATA 1006
 Db 1961 CCATCAAAATATCAATCAATCTTCTCAAGAGATTTAG--ACAGAACAATTTGCAAAATTCATC 2019

QY 1007 TGAACCAACCAACCAAAAAAAAAAAAAAAAACCCGATAGCCAAAGCAAGCTTAGCAAAA 1066
 Db 2020 TGAATTAAC-----AAAAAAAAAGAGATAGCTTAAGCTATCTTCAACATA 2065
 QY 1067 AGAACAAATCTGAGAGCATCATATTACCATTTCAAACTATACTACAAAGCTATATCA 1126
 Db 2066 AAGAGACTTCAGGGGAAATCACTATCTCCCTGAACCTCAACAGATATTAAGACAAATAGTA 2125
 QY 1127 CCAAAACATGAGCACTGACATTAATATAGCAATAGCAATAGCAAGAAAGAGAGA 1186
 Db 2126 TAAAACTGATGTATGTGTACAGAGACAGACATAGAACCAATGAAATGAATTGAG 2185
 QY 1187 ATCCGAATTAATAGCAAAATTAATTAATAGCAACTGATTTTTCAGAAACCAAAACA 1246
 Db 2186 ACCGAATTAAGAACCAACACACTTATGTGTCACTTATTTTGCAGAAAGCAAAACCA 2245
 QY 1247 TAAAGTGGGAAAAGACATTTCTAGTTAACAAATGCTGAGATTAATGCAAGCCAT 1306
 Db 2246 TCCAAATGAAAAAAGATAGCATTTTCAAGCAATGCTGCTGTTCAACTGAGAGCCAT 2305
 QY 1307 GTGAAAGATGAAATCTGATCCCTTGTCTCACTTAATACAAATTTGATCAAGATG 1366
 Db 2306 GTAGAAAGATGCAAGTCAATCATCT--TATCACTCTGACAAACCTTAAGTCAAGTGG 2364
 QY 1367 ATCAAAAGCTTAATCTGAGACCTTAACCAATTAATAATCTAGAAATAGATACAGAAA 1426
 Db 2365 ATCAAGAGACTTCACATCAACCAACACACACTCAATTAATGAGAAACCACTAGGAG 2424
 QY 1427 ATGCTTACATTCATCTAGGCAAGACTTATGAGCCAGAACCCAAAGTAATGCA 1486
 Db 2425 CATCTGGAACACATGAGGCACTGAAAAAATTTCTGGAACAAACCAATGAGCTTATCT 2484
 QY 1487 ACAAACCAAAATTAATGATAGGCTTAATTAATTAAGCTTTTGGCAGCAAAA 1546
 Db 2485 CTAGATCAAGATGACAAATGGAATCTATTAATCTGCAAGCTTCTGTAAGGCAAG 2544
 QY 1547 ACATCATTAAGCAGAGCAACAGAACCCAGAGTGAAGAAAAATCTTCAAACTTA 1606
 Db 2545 GACATGCTGTATAGAGCAAAACGGAACCAACAGATTTGGAAAGATCTTTACCAATCT 2604
 QY 1607 GCATCTGACTAAGACTAATATCCGATCCACAGAACTCAACCAATCAATGACAGAG 1666
 Db 2605 ACAACAGATAGAGGCTTATTTTCAAAATATTAAGAACTCAAGAGTTAGACCGCAGG 2664
 QY 1667 AAAGCAAAATCCCAATGAAAGAGTGGCTAAGACATGAATGACATTTCCAAAAGAA 1726
 Db 2665 GAAACAAATTAACCTTAATTAATAATGAGGCTCAGAGCTTAACAAAGATTCACAGCTAG 2724
 QY 1727 GATATCAATAGGCAACAAACA--GSAAAAAATGCTTAACATCACTAATGTTAGGAA 1784
 Db 2725 GATGCGGAATGGCTGAGAAACACCTTAAGAAATGTCAACTTTTGTGTAAAGGAA 2784
 QY 1785 ATGTAATTAACACTGTAAATGCGATACCACTTACTCTGCAAGATGCTAATTTTA 1844
 Db 2785 ATGCAATCAAAACAACCTTGAGATTTTCACTCAACAGCTGCAATGCTTAAGTCAA 2844
 QY 1845 AATCTAATAATAATGATGTTGGGTGCTGAGTGAATGAAGAACTTTTACACTGC 1904
 Db 2845 AACTGAGTGAACGAGATGCTGGCGAGATGAGAAAGAGAACTCTCTCATTTGT 2904
 QY 1905 TGGTGGAAATGTAACTTTCGCAACCACTATGAAAAAGTGTGAAATTTCTTAAGAA 1964
 Db 2905 TGGTGGAAATGTGAAGATGTTAAACATCTGTGAATATAGTGTGAAGGTCTCTCAGAAA 2964
 QY 1965 CTAAAGATGATGACCATTTTATTCAGCAATCC 1999
 Db 2965 TTGACATTTGAACCTGCTAGAGATCCAGCTATACC 2999

RESULT 5
 AY539883
 LOCUS

AY539883

4071 bp

mRNA

linear

HTC 21-MAR-2004

DEFINITION Rattus norvegicus LRG00132 mRNA, complete cds.
 ACCESSION AY539883
 VERSION AY539883.1 GI:45478103
 KEYWORDS HTC.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 4071)
 Xu, C.S., Zhang, L., Chang, C.F., Han, H.P., Wang, G.P., Chai, L.Q.,
 Yuan, J.Y., Yang, K.J., Zhao, L.F., Ma, H., Wang, L., Wang, S.F.,
 Xiang, X.K., Shen, G.M., Shi, J.B., Rahman, S., Wang, Q.N. and Zhang, J.B.
 Liver regeneration after PH
 Unpublished
 2 (bases 1 to 4071)
 Xu, C.S., Zhang, L., Chang, C.F., Han, H.P., Wang, G.P., Chai, L.Q.,
 Yuan, J.Y., Yang, K.J., Zhao, L.F., Ma, H., Wang, L., Wang, S.F.,
 Xiang, X.K., Shen, G.M., Shi, J.B., Rahman, S., Wang, Q.N. and Zhang, J.B.
 Direct Submission
 Submitted (03-FEB-2004) Henan Bioengineering Key Lab, Henan Normal
 University, NO.148 Jianshe Road, Xinxiang City, Henan 453002,
 P.R.China
 FEATURES
 source
 1..4071
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 1..4071
 /note="liver regeneration-related"
 /codon_start=1
 /product="LRG00132"
 /protein_id="AA566223.1"
 /db_xref="GI:45478104"
 /translation="MPTYTKLNNALLNDNLVKEIKIKIKDFLENEENEGTYPNLM
 DTKAVLRGKFEISACRKKQRAYVSLAHKALQKQKAPRPSRRREIILKRAE
 FNOIEAVIKGLPTKSPGDSABEYPTIEDLIPILSKLTHKIETDALSPTIE
 ATTLLPKPHKDTTKKRNFRPISLMDAILKILKILNRIQEHKTIYDQGFIPG
 MOGFNIRKTIINVIHYINKLEQNHMIISDAKAFDKIHPMKIVLERIGIQQPY
 NIEIKIRAEINOVETKRTIERINRTKSPFEDKINKDKDLATGHRGCVIOGKLR
 NEKGIITDDEEIOKIRSYKMLSTKFNLOEMDFLRVYVSKLNOGIONLNP
 IPKEIEAVIKGLPTKSPGDSABEYPTIEDLIPILSKLTHKIETDALSPTIE
 EATITLIPKHKDTTKKRNFRPISLMDAILKILKILNRIQEHKTIYDQGFIP
 GNOGFNIRKTIINVIHYINKLEQNHMIISDAKAFDKIHPMKIVLERIGIQQPY
 LANVKAIVSKPVANIKLNGEKLEAPIKSGTSCPSPLFNIVLELRAIQOGE
 IKGIQIKGEKVKISLPADMIIVILSDPSSTRELLKILNFSVAGKIKINSNVAF
 YTEKQAEKIEKRETPPIDPNNIKVLGVLTQVODLVKQDLYKPTLKEIEEDLRK
 DLPSGWRINMFMALIPKAIYRENAIPKIPLOPKELDPTICRPIINMKPRILAK
 ALINMKTSGGITIPELKQYRYAIVIKTAMTYRDRDIDQWRIEDPENAPHYGHLI
 FDGAKTIQWKDQISFKWCFWNRATCRWQIDPSLSPCTKLSKRIKDLHRTPL
 KILBEKLGKLEHMGKPLNKTPTMAVLRSLIDKDLIKQSFCADQIVATKRO
 PTDMEKIFTPPTDRLGILSKIVKELKALCRFTNNPIKMGSELNKEFTAKEPMAR
 HLKKGASLVIREMOIKTLRFLPTPLAKIKNSGDSKMGSGEGTLLHCPMARP
 LVPEFWKSWVRFLKLDILBEDPAIPLGIVPKDASTYRDRCTSMETALPLTIAR
 KMERPRSTREMIQKMTYITMYEYSLIKNNEFKYGVKLEGNITLSLSTOSQDI
 YGEGSFWLTPVDVSHGKIKIACSPISNAPLKITFTINAPMGENISVIRKAGD
 LRQDMVLQIIVQMDNVMLQEGDMQIIVGICALPAQFIEMVDPVATLAKIHLHS
 GLIGPKENTIKKMFQSHNLKEDYEKKPQSRDEAF"

ORIGIN
 Query Match 9.3%; Score 801.8; DB 3; Length 4071;
 Best Local Similarity 65.6%; Pred. No. 1,56-167;
 Matches 1321; Conservative 0; Mismatches 662; Indels 32; Gaps 9;

QY 1 TGTATGAGCGCAATGCTTTAATACCAAAACAGAGAAAGATATA--CAAAAAGAAA 58
 DB 1346 TCTACAGAGCCAACTTACTCTTATCTTAAACCAACAAAGACACAAAGAAAGAGA 1405
 QY 59 ACTATGAGCCAGTACCACTGATGATATATACATGACGAATCCCCAACAAATATCTAGCTA 118
 DB 1406 ACTTCAGACCAATTTCCCTTATGAAATATGACGCAAAAATATCTCAATAAATCTGGCAA 1465
 QY 119 ACCCAATCCAGACGATATCAAGAGATATCAACCACTGTCAAGTGGGTTTCATACCAAG 178

DB 1466 ACCCAATCCAGACGATCAAAAACATCAATCCACCAATGATCAAGTAGGCTTCATCCAG 1525
 QY 179 GGGTGA--GGATAGGTTAATACATACAAAGTCAATTAATGTGATACATACATAAAGAAA 237
 DB 1526 GCATGACGGATGGTTTAAATATACGAAAAACATCAACGGATGCATTAATATAACAAAC 1585
 QY 238 TTAATAAACAATAATCACATGATCATCTCAATAGTGTGAAAAAGCATTTGACAAATCT 297
 DB 1586 TGAAGAACAAGAACCAATGATCATCTCAATAGTGTGAAAAAGCATTTGACAAATATTC 1645
 QY 298 AACATTTCTTATGATTAATAAACCCTTGACAAAATCGACATAGAAAGACATACCTTAATG 357
 DB 1646 AACACCCCTTCATGATTAATAAGTCTTGAAAGAAATAGAAATCAAGGCCATACCTTAACA 1705
 QY 358 TAAATAAAGCCATATATATGACGACCCACAGCAAACTTATATCTGAATGGGAAAGTTGA 417
 DB 1706 TAGTAAAGCCATATATACAGCAAAACGATGCTTAACATTAATGAGAGAAACTTG 1765
 QY 418 AACATTTGCTCTGAGAACTGGAACAAGACAGAGATG--CTACTTCAACACTTCTATTCA 476
 DB 1766 AACCAATCCCATTAATATCAGGACCTAGACAGGCTGCCACATCTCTCCCTACTTATTCA 1825
 QY 477 ACATAGTAGTGAAGTTTATGCGAGGATCATGACAGAGAAAGAAAGAAATGAAAGGCACCC 536
 DB 1826 ATATAGTTCTTGAAATCTAGCCAGACATCAAGACAAAGAAAGATCAAAAGGATAC 1885
 QY 537 AAATCAATAAAGAGAGAGTCAAACTGCTCCCTGTTCACTGATGATGATGTATACCTAG 596
 DB 1886 AGATCGAAAGAAAGAGAGTCAAAATATCATATTGAGATGATGATATATATTTAA 1945
 QY 597 AAAACCTTAAGACTCATCAAGAAAGCTCTAGAACTGATACATTAATTCAGTAAAGTTT 656
 DB 1946 GTGATCCCAAAAGTTCCACAGAGAACTAATAAGCTGATTAACAATTCAGCAAAAGTGG 2005
 QY 657 CAGATTCAACTTAATGTCACAAATCACTAGCACTGCTATACCAACAGTGAACAAG 716
 DB 2006 CTGGATATTAATAATTAATCAATTAATCAATGATGCTCTCTTATACAAAGAGAAAG 2065
 QY 717 CTGAGATCAATCAAGAACTCAAACTTTTCAATAGCTGTAA-----AAATACT 770
 DB 2066 CCGAGAAAGAAATTAGGAGAAACGACCTTCTATATAGACCCAAATATATTAAGTACC 2125
 QY 771 TAAGAATATTCTTACCCAGAGAGTGAAGACCTTCAAGAGAAACTCAAAACACAG-- 829
 DB 2126 TCGGTGACTTAAACCAAGCAAGTAAAGATGTGACATTAAGAACTTCAAGACACTGA 2185
 QY 830 ----CTGACATCATATGATGACAAACAAAGTGAACACATCCCATGCTCATGTGGGTA 886
 DB 2186 GGAAGAAATTAAGAAAGACCTGAGAAAGATGGAAGATCTCCCATGCTCATGTGGGCA 2245
 QY 887 GAATCAATATTGTAATAAATGACATATATGCAAAAGCAATCTCAAGTTCAATCAATTC 946
 DB 2246 GGAATTAATATGTAATAAATGACATATATGCAAAAGCAATCTCAAGTTCAATCAATTC 2305
 QY 947 CCAACAAATATGATCATCATCTTCTTCAAGAACTAGAAAAAAACAATCTTAAATTCATA 1006
 DB 2306 CCAATCAAAATATCAATCAATCTTCAAGAGTTAG--ACAGAACAATTTGCCAATTCATC 2364
 QY 1007 TGAACACACACCAAAAAAAAAAAAAAAAAAAAAACCGCATAGCCAAAGCAAGCTTAGCAAAA 1066
 DB 2365 TGAATAATAC-----AAAAAACCGAGATAGCTTAAAGCTATCTTCAACATATA 2410
 QY 1067 AGAACAATCTGAGAGGATCATCAATTAACCTTTCAACTATATCTACAGAGCTTAATCA 1126
 DB 2411 AAAGACTTCAGGGGAATATCATATCTTGAACCTCAAGCAGTATATACAGAGCAATATGA 2470
 QY 1127 CAAAACATCATGACCTGACATTAACATTAAGCAATGACATGACAAATGAAAAAGAGAGA 1186
 DB 2471 TAAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2530
 QY 1187 ATCCAGAAATAAAGCCAAATTAATTAAGCAACTGATTTTGGCAAGCAAAAGAAACA 1246

Db	2531	ACCAGAAATGACCACACACCTATGCTGACTGATTTTGGACAAAGAGCCAAAACA	2590
Qy	1247	TAAAGTGGGGAAAAAGACATTCTAGTTAAACAATGCTGTGAGATTATTGGACACCACAT	1306
Db	2591	TCCAATGGAAAAAAGATGACATTTTTCAGCAATGGTGTGTTCACTGGAGGGCAACAT	2650
Qy	1307	GTGGAGAAATGAAACTGGATCCCTGTCTCACTTAATTAACAATAATTGAATACAGATGG	1366
Db	2651	GTAGAGAAATCAGATCGATTCATCTCT-TATCACCCTGTACAAAGCTTAAGTCCAAATGG	2709
Qy	1367	ATCAAGAAGCTTAAATCTGAGACCTTAAACCATATAAAATTTCTAGAAGATPACATGAGAAA	1426
Db	2710	ATCAAGAGACCTCCACATCAAAACGAGACACACTCAAACTTAATAGAAAGAAAACTAGGGAAG	2769
Qy	1427	ATGCTTTCTAGACATTCACCTTAGGCAAGACATTCATGCGCCAGAACCCAAAAGTAAATGCA	1486
Db	2770	CATCTGGAGCACTGGGCGACTGGAAAAATTTCTCGAACAAAACCCAAATGGCTTAATGCT	2829
Qy	1487	ACAAAAAACAAAATTAATATGATATAGGACTTAATTAATCTAAAAAGCTTTGGCGACGAAA	1546
Db	2830	CTTAAGATCAAGATTCGACAAATGGGATCTCAATAAACTCGAAAGTTCTGTAGAGCAAG	2889
Qy	1547	ACAAATCATTTAGCAGAGCAAAAGACAAACCCACCGAGTGAAGAAAAATCTTCACAAACTTA	1606
Db	2890	GACACTGTGTTAGACAAAAACGGACACACAGATTGGGAAAAGATCTTTACCATCTCT	2949
Qy	1607	GCATCTGACTAAGGACTAATATCTCGGAATCCACAAAGAACTCCAAACAAATCAGCAAGAG	1666
Db	2950	ACAAACAGATPAGAGGCTTATTTCCAAAATATACAAAGAACTCAAGAAATGATATGCCGACAG	3009
Qy	1667	AAAGCAAAACATCCCATTAAGATAGAGGGGCTAAGGACATGAATATGACATTCGCAAAAAGAA	1726
Db	3010	GAAACAAATPAACTCCATTATTAATAAAATGGGGTTCAAGGCTPAAACAAAGAAATTCACAGCTAAG	3069
Qy	1727	GATATACAAATGGCCCAACAAACA--GGAATAAAATCTTAACTACATCACTATGATTAGGAAA	1784
Db	3070	GAAATGCCAAATGGCGGAGAAACACTTAAGAATTTTCAGCATCTTTAGTCATAAGGAAA	3129
Qy	1785	ATGTAAATCAACACTGTAAATGCGATACACCTTACTCTCGACAGAAATGCTCAATATTTAA	1844
Db	3130	ATGCAAAATCAAAAACAACCCCTGAGATTTTCACTTCAACACAGATGGCGAATGGCTAAGATCAAA	3189
Qy	1845	AAATCTAAAAATTAATATAGATGTTGTGGGTCTGTGGTGTAAAGAAACCTTTTCACTGCG	1904
Db	3190	AACTCAGGTGACAGCAGATGCTGGCGAGAGATGTGAAGAAAGGAAACCTCCTCCATTTGT	3249
Qy	1905	TGTTGGGAATGTAAACTCTTGCGCAACCACTATGAGAAAAAGTGTGAAAAATTTCTTAAGGAA	1964
Db	3250	TGGTGGGATTTGCAGACTGCTGTAAACCAATTTCTGGAATCAGTCTGAGAGTTCTCTCAGAAA	3309
Qy	1965	CTAAAAAGTATGATGACCAATTTTGATCAGCAATCCC	1999
Db	3310	TTGACATTTGAACCTGCTGAGAGATCCAGCTATACC	3344
RESULT 6			
LOCUS	AY539886	4356 bp	mRNA linear
DEFINITION	Rattus norvegicus LRRG00135 mRNA, complete cds.		HTC 21-MAR-2004
ACCESSION	AY539886		
VERSION	AY539886.1	GI:45478109	
KEYWORDS	HTC.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		
	Rattus.		
REFERENCE	1 (bases 1 to 4356)		
AUTHORS	Xu,C.S., Zhang,L., Chang,C.F., Han,H.P., Wang,G.P., Chai,L.Q.,		
	Yuan,J.Y., Yang,K.J., Zhao,L.F., Ma,H., Wang,L., Wang,S.F.,		
	King,X.K., Shen,G.M., Shi,J.B., Rahman,S., Wang,Q.N. and Zhang,J.B		
TITLE	Liver regeneration after PH		
JOURNAL	Unpublished		

REFERENCE	AUTHORS	TITLE	JOURNAL
2 (bases 1 to 4356)	Xu, C.S., Zhang, L., Chang, C.F., Han, H.P., Wang, G.P., Chai, L.Q., Yuan, J.Y., Yang, K.J., Zhao, L.F., Ma, H., Wang, L., Wang, S.F., Xing, X.K., Shen, G.M., Shi, J.B., Rahman, S., Wang, Q.N. and Zhang, J.B	Submitted (03-FEB-2004)	Henan Bioengineering Key Lab, Henan Normal University, NO.148 Jianshe Road, Xinxiang City, Henan 453002, P.R.China
FEATURES	Source	Location/Qualifiers	
		1..4356	
		/organism="Rattus norvegicus"	
		/mol_type="mRNA"	
		/db_xref="taxon:10116"	
		1..4356	
		/note="1 liver regeneration-related"	
		/codon_start=1	
		/product="LRG00135"	
		/protein_id="AA566226.1"	
		/db_xref="gi:45478110"	
		/translation="MNITGNNHYSLSLININGLNSPIKRRHLTMNICDEPAFCCLD ETHLADKRHYRYKNGKWTQTTPQANGQCKQAGVALISNKPQLKYIKDKGHRIFP KKHIDHDELSTINITYAPVTRAPVITYIKETLKLKTHIAPHTIIVGDPNPLSSMDSW QKNSDYLRLREWSQMDLTDIYRTFYPKAGQTFPSAPGQFSKIDLIIGKTLNGLN YKRIELIICVLSDHGLGLVFNPNKNGMPYTKNALINDLVYEEIKKIKNPLNGLN FHNENETTPNLMQDMTKAVLYSGKILALSACRKOERYALVSLHAKLQGEATP RRBOEIIKLAELINOVETKTRTIRIRTSWEPKINKIDEXLALTEGHECOYIN KIRNEGDDITTDSEITQKILSYTKNYKSTKFNLOEMDNFDRYVSKLNOBOINO NNPITPEKILBIVIKGLPTKYSPPGDFSAEYVLEDFLPILISIKFHIETDGLQV SEVEATITLIPKSHDPTTKENFRPISLNIIDAKLNLKANIISNKHITTHIHQOVL FTMGQGMENIKRTKNWVHYINKLEQNHNIISLDAEKAFDILQHPFMVLEKRGIL GPLINVRAIYKSPVANKLNGEGLQALPKSGRQCCPLSPFNIVLEVARIAIRI QKEVGIOIGDEQVYKISLPADMDIAYLSDPSRRELKILNNFIVAAQYKNSKYS ALYYTEKQAEKRIETPTPIIDNNIKYIGULTYQVYDLVYKNEKTRKEIEEDLI RMDVPCSWGIRINIVKPAIIPKATVFNMIPIKIPQFPEKLDITICFTNNKKP IKALINKRTSGGTTIPELQYVRAIVITKAWYARDQIDQWNIIEDENNPHYV HLIFDGAQTIOMKDQISFSKQCNFRARCTRWQIDPCLSPTCKSKWIKDHLIK DILKILIEELGGRHEMGTKNFKLTPMVALRSRIDKDLIKQSPCKADQTVR KQOPDMWEIKFTNPPTDGLISKIYKELKLDRETNNP1KKKGSILNKPFAEBCR AEHLKKGSTSLIRENOIKTTLRPHLTPVRLAKIKNSGDSRCMRCGGRGTLHCPM DCLVLPFWKSYWRFLRKLDLEPEDPAIPULGIYFADASTKRDTCSTWFLAALPT ARWKEPRCPSTBEWIIQMWYIYTMETYSALIKNEFWKPEGVKLELENIILSETQSS KIDHDLQTCILMEHCIDQCMWGGSDDHCAQMAVAYDQITDRTNLTCLPHSSTYV LAKPOGECAMNOCGRKGDELGLGALAGVNRCDPLQOCQALKEBRLLAESWV IRCPAGPHVMELEWGNPPPPPHRPGSDPCVTRPVPEGRKPFVGLLLGMVGQ STAAEQDPECCLATY"	
ORIGIN			
Query Match	9.2%; Score 790.6; DB 3; Length 4356;		
Best Local Similarity	65.2%; Pred. No. 4,9e-165;		
Matches 1314; Conservative	0; Mismatches 669; Indels 32; Gaps 9		
1	TGTATGAGCCAAATGTCACTTTAATACCAAAAACGAGAAAGATATTA--CAAAAAAGAA	58	
1529	TCTACGAAGCCAAATTACTCTTAATACCTAATATCAACAAGAACAAAAAGAGA	158	
59	ACTATGACCAAGTACCACTGATGATATATACATGACGAAATTCCTCCAAATAATCTAGCTA	118	
1589	ACTTCGACCAATTTCCCTTAATGAAATATCGACGAAAAAATATCTCAATAAATAATTCG3CA	164	
119	ACCAATCCAAAGCATATCAAGAAGATTAATTCACACATGTGCAAGTGGGTTTCATACAG	178	
1649	ACCAATTCAGAGCATTCAAAAACAATCATCTACCATGACCAAGTAGGCTTATCATCAG	170	
179	GGGTGCA--GGATAGTTAAACATACACAAGATCAATTAATGTGATACATACATTAACGAA	237	
1709	GCATGACGAGATGTGTTAATATAGGAAAAACAATGAAGCGTTCATTATATTAACAAT	176	
238	TTAAAAACAAAATTCATCATGATCATCTCAATAGATGCTGAAAAAGCTTTGACAAATCT	297	
1769	TGAAAGAACAAAACCATGATCATTTTATTAAGCGCTGAAAAAGCATTTGACATTAATTC	182	
298	AACATTTCTTAAGTATAAACCTTCACGAAAATCGACATGAAAGACATACCTTAATG	357	

Dp	1829	AACACCCCTTATGATTAAGAATGCTCGAAGAAAGATTGGAATTCAGAGCCCATCTGAAACA	1888
Qy	358	TAAATAAAAGCAATATATGACGAGCCCAACAGCAAACTTATATCTGAATGGGAAAAGTTGA	417
Dp	1889	TAGTAAAAAGCCATATACAGCAAAACCAAGTTGGCTAATATTAATCTAAATGGAAGAAAATCTC	1948
Qy	418	AAACATTTGCTCCTGGAAGCTGGAACAAACAAGAAATGCTACTTTTACACATTCATATCA	476
Dp	1949	AAGCAATCCCACTAAATACAGGAACTTAACAAGGTGCGCACTCTCTCTCACTTATATCA	2008
Qy	477	ACATAGTAGTGAAGTTTATAGCCAGAGAACTCAGACAAGAAAGAAATCAAGGGCACCC	536
Dp	2009	ATATATGTTCTTGAAGTTCTAGCAGAGCAATCAAGAAAACAAAGAGGTCAAGGGGATAC	2068
Qy	537	AAATCAATTAAGAAGAAAGTCAAACTGCTCTCTGTTCACTGATATATGATTTATACCTAG	596
Dp	2069	AGATCGGAAAAAGATGAAGTCAAAATATCACTGTTTGACAGATGATGATGATCAATTTTAA	2128
Qy	597	AAACCCCTAAAGACTCATCCAGAAAGCTCCAGAACTGATATCATTAATTCAGTAAAGTTT	656
Dp	2129	GTGATCCCAAAAGTTCCACAGAGAACTATCAAGCTGATTAACAACCTTCATCAAAAGGG	2188
Qy	657	CAGGATACAAACTAAATATACAAATATAGTAGCACTGCTATACCAAGTGCACAG	716
Dp	2189	CTGGGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2248
Qy	717	CTGAGAACTCAATCAAGAACTCAAACTTTTACATAGCTGTAA-----AAATACT	770
Dp	2249	CCGAGAAAGAAATTAAGGAAAGACACCTTCAATATAGACCAAAATTAATTAAGTACC	2308
Qy	771	TAAGAATTTCTTACCCAAAGAGTGAAGAGCTCTACAGAGAAATCTACAAACACAG-	829
Dp	2309	TCGGTGTGACTTTTAAACGAAGATGAAGATCTGTACAAATGAAGACTTCAAGACACTGA	2368
Qy	830	---CTGACATCATATGATGACACAACAAGTGGAAAACATCCCAATGCTCATGAGTGGTGA	886
Dp	2369	GGAAGAAATTTGAAGAAAGACCTCAGAGATGGAAGATCTCCACATGCTCAAGATTGGCA	2428
Qy	887	GAATCAATATTTGTGAAAAATGAACATATTTGCCAAAAGCAATCTACAGTTCAATGCAATTC	946
Dp	2429	GGATTAATATATGTAATAAATATGGCCATTTTACAAAAGCAATTAACAGATTCATGCAATCC	2488
Qy	947	CCACCAAAATATCATCATCATTTCTTACAGAACTGAAAAAAACAATCTAAATTCATA	1006
Dp	2489	CCATCAAAATTCCAATCCAAATTTCTTCAAGAGTTAG-ACAGAACAAATTTGCAAAATTCATC	2547
Qy	1007	TGGAACAACAACCAAAAAAABAAAAAABAAAAAABAAAAAABAAAAAABAAAAAABAAAAA	1066
Dp	2548	TGGAATTAAC-----AAAAAACCCAGATATGCTTAAGGCTATCTTCAACATATA	2593
Qy	1067	AGAACAAATCTGAGGAGCATCATATTCACCATCTTCAAACTATACAGAGCTATTAATCA	1126
Dp	2594	AAAGGACTTTCAGGGGGAACTCACTATCTCTGAACCTAAGCAATTAACAGACATAGTGA	2653
Qy	1127	CCAAAACTATGCGCATGACATTAABACTAGGCACTAGACCAATGAAAAAGAGAGA	1186
Dp	2654	TAAAACTGCAATGATATTTGTAACAGAGACAGACAGATAGACCAATGGAATGGAATTGAAG	2713
Qy	1187	ATCCAGAAATTAAGCCAAATTAATTAAGCCAACTGATTTTGAACAAGCAACAAAAACA	1246
Dp	2714	ACCCAGAAATGAACCCACACACCTATATGCTCATTTGATTTTGAACAAGAGACCAAAACA	2773
Qy	1247	TAAATGGGGGAAAAAGCACTTCTAGTTAAACAATGGTGCTGAGATTATTTGGCAACCAAT	1306
Dp	2774	TCCAAATGGAABAAAAAGATAGCATTTTTCAGCAAGTGTGCTGTTCAACTGGAGGGCAACAT	2833
Qy	1307	GTGGAAGATGAACCTGAGTCCCTGTCTCTCACTTAATACAAAAATTTGATACAAATGG	1366
Dp	2834	GTAGAAGATGCAATGATGATCAATGCT-TATCACCTGTATCAAAAGCTTAAGTCAAGTGG	2892
Qy	1367	ATCAAAAGACTTAATCTGAGACTTAABAAACATAAAAAATCTAGAGAGTACATCAAGAAA	1426
Dp	2893	ATCAAGAGCTTCCACATCAACACAGACACTCAAACTTAATTAAGAGAAAACTAGGAGG	2952

QY	1427	ATGCTTCAGACATTCACCTTAGGCAAGAAAGCTTCACATGCGCAAGAACCCAAAGATAATGCA	1488
Db	2953	CATNAGGGAACACATGGGCACTGGAAAAAATTTCTCGAACAAACACCAATGGCTTAATGCT	3012
QY	1487	ACAAAAACAAAATTAATATAGATATAGAGACTTAATTAATCTAAAAAGCTTTTGGCAGCAAAA	1548
Db	3013	CTAAGATCBAAGATCGACAAATGGAGTCTCATTAATACATCGCAAACTCTGTAAAGGCAAG	3072
QY	1547	ACAATCATTAGACAGACAAACAGAACCAACCGAGTGAAGAAAAATCTTCAAACTAA	1608
Db	3073	GACACTGTGGTTTAGGACAAAACGGCAACCAACATTTGGAAAAAGATCTTTACCAATCCT	3132
QY	1607	GCATCTGCTAGGACCTAATTCGGGAATCCAGAAAGAACTCAAAACAAATCAGCAAGAG	1666
Db	3133	ACAAACAGTAGAGGCTTAATATCCAAAATATACAAAGAACTCAAGAACTTACGCCGACGG	3192
QY	1667	AAAGCAAAACAATCCCATGAAAGAGTGGGCTAAGACATGATATAGACAATTTCTCAAAAGAA	1728
Db	3193	GAACCAATTAATCCCTATTAAAAAATGGGGTTCAAGCTAAACAAAGAAATTCACAGCTAG	3252
QY	1727	GATATACAAATGGCCACAACAAACA--GGAAAAATGCTTAACATCATCAATATAGGGAA	1784
Db	3253	GAATGCCAAATGGCTGAGAAAACACTAAGAAATGTTCAACATCTTTAATCATAAAGGAA	3312
QY	1785	ATGTAAATCAACACTGTAAATGCCATACCACTTACTCTCGAAGAAATGCTATATTTAA	1844
Db	3313	ATGCAAAATCAAAACAAACCTCTGAGATTTCACTTCAACACGACGTCATTTGGCTAAGATCAA	3372
QY	1845	AAATCTAAAAATTAATATAGATGTTTGGTGGGCTGTGTGATTAAGAAACACTTTTACACTGC	1904
Db	3373	AACTCAGGTGACAGCAAGATGCTGGCGAGAGATGTGGAAGAAAGAGAAACATCTCTCATTTGT	3432
QY	1905	TGTTGGGAATGTAAACTTTGGCGAAACAATAAGAAAAACAGTGTGAAATTTCTTAAGGAA	1964
Db	3433	TGGTGGGATTCACAGCTGTGTAAACCAATTCGAAATCAAGTCTGAGGTTCTCAGAAAA	3492
QY	1965	CTAAAGATGATCGACCATTTGATTCAGCAATCCC	1999
Db	3493	TTGACATTTGAATGCTCGAGAGATTCAGCTAATACC	3527

RESULT 7				
LOCUS	BC038425			
DEFINITION	BC038425 Homo sapiens, clone IMAGE:4818574, mRNA.	2488 bp	mRNA	linear
ACCESSION	BC038425			
VERSION	BC038425.1	GI:23512256		
KEYWORDS	HTC.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 2488) Strausberg,R. Direct Submission			
REFERENCE	Submitted (01-OCT-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA			
AUTHORS				
TITLE				
JOURNAL				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov			
COMMENT	Contact: MGC help desk Email: ggapbs-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amg@bcm.tmc.edu Gunnarste, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,			

Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLIN at: <http://image.bnl.gov>
Series: IPAX Plate: 71 Row: a Column: 12
This clone has the following problem: retained intron.

FEATURES

SOURCE

1..2488
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4818574"
/tissue_type="brain, hippocampus"
/clone_lib="NIH MGC_95"
/lab_host="DH10B"
/note="Vector: pBluescript"

ORIGIN

Query Match 8.9%; Score 770; DB 3; Length 2488;
Best Local Similarity 83.7%; Pred. No. 1.8e-160;
Matches 948; Conservative 0; Mismatches 165; Indels 19; Gaps 6;

```

Qy 872 GCTCATGGATGGGATGAATCATATTTGTAAGATGACATATTGCGAAAGCAATCTAGA 931
Db 1325 GCTCATGGATGGGATGAATCATATTTGTAAGATGACATATTGCGTAAGCAATCTAGA 1384
Qy 932 AGTTCAATGCAATTTCCACCAAAATATCATCATCTTCTTCACAGAACTGAGAAAAACA 991
Db 1385 AATTCACACAAATTTCCACCAAAATATCATCATCTTCTTCACAGAAATAG-AGAAAACA 1443
Qy 992 ATTCTAAATTCATATGAGAACCAACCAAAAAAACCCTATGCGCAAG 1051
Db 1444 ATTATTAATTCATATGAGAAC-----CAAAAAAGGCCCATATCCAAAG 1489
Qy 1052 CAAGCTTAGCAAAAGAAACAATCTGAGGATCATATACCCTTCAAACTATCT 1111
Db 1490 CAATCTAGCAAAAGAAACAATCTGAGGATCATATCTGATTTCAAACTATGCT 1549
Qy 1112 ACAAGCTATATATCAACCAAAATCATGSCATCATATAAATAGGACATAGACCAAT 1171
Db 1550 ATTAAGGCTATATGTCACCAAAAGCATGCTATGATATAAACAAGGACATAGACCAAT 1609
Qy 1172 GAAAAAGAGAGAAATCCAGAAATTAAGCCAAATTAATAGCCAACTGATTTTGACA 1231
Db 1610 GGAACAGAAAGAGAAATCCAGAAATTAAGCCAAATTAATAGCCAACTGATTTTGACA 1669
Qy 1232 AAGCAAAACAAAACATTAAGTGGGAAAGACATTTAGTTAAACAATGGTCTGAGATT 1291
Db 1670 AAGCAAAACAAAACATTAAGTGGGAAAGACATTTAGTTAAACAATGGTCTGAGATT 1729
Qy 1292 ATTGGCAAGCCATATGTAAGATGAATGAATGATCCCTTGTCTCATCTTAATACAAA 1351
Db 1730 ATTGGCTAGCCATATGTAAGATGAATGAATGATCCCTTGTCTCATCTTAATACAAA 1788
Qy 1352 ATTGATTAAGATGATCAAGACTTAATCTGAGGCTTAATCCCTTAATTAATCTAGAA 1411
Db 1789 ATCAATCTAAGATGATTAAGGCTT-AACTTAAGGCTTAATCCCTTAATTAATCTAGAA 1847
Qy 1412 GATAACATCAGAAAAATGCTTTAGACATTTCACTTGAAGCAAACTTCAAGCAAGAAC 1471
Db 1848 GATAACATTTGGCAAAACCTTTAGACATTTGCTTGAAGCAAGATTTCAATCAAGAAAC 1907
Qy 1472 CCAAAAGTAATGCAACAAAACCAAAATTAATATGATGAGCTTAATTAATCAATCAAG 1531
Db 1908 CTAAGAGCAATATGCAATTAATCAAAAGATTAAT-GCTGGGACTTAATTAATCAAGAGC 1966
Qy 1532 TTTTGGGAGCAAAACATCAATTAAGCAAGCAAAAGCAACCCAGCGAGTAGAGAAA 1591
Db 1967 TTTTCTATGCAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 2026
Qy 1592 ATCTTCAACCAATGATCTGATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1651

```

```

Db 2027 ACCTTCACATCTATATGCTGACAAAGACTAATATATCCAGATCTATATAGAGCTCAA 2086
Qy 1652 CAATTCAGCAAGAAAGCAAAACAATCCATGAAAGTGGGCTTAAGACATGAAATGA 1711
Db 2087 CAAATTTAGCAAGAA-AAAAAAATATATCCATCAAAAGTGGGCTTAAGACATGAAATGA 2145
Qy 1712 CAATTCCAAAAGAGATATATCAAAATGCGCAACAAACAGAAAAAATGCTTAACATCACT 1771
Db 2146 CAATTCCAAAAGAGATATATCAAAATGCGCAACAAACATTAATAAATCCCAATCACT 2205
Qy 1772 AATGATTAAGCAAAATGTAATATCAACATGTAATGCGATPACACTTACTCTGCAAGAT 1831
Db 2206 AATGATTAAGCAAAATGTAATATCAAAACCAATGATGATACACTTACTCTGCAAGAT 2265
Qy 1832 GGTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1891
Db 2266 GGCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2325
Qy 1892 ACTTTTACACTGCTGGTGGGAAATGTAATCTTGCGCAACCACTATGAAAAAGTGTGAA 1951
Db 2326 ACTTCAACATGCTGCTGGTGGGAAATGTAATCTATATAGCACTATGAAAAAGTGTGAG 2385
Qy 1952 ATTTCTTAAGAACTTAATAAGTAAATGCAATTTGATTCAGCAATCCCAATTA 2003
Db 2386 ATTCCTTAAGAACTTAATAAGTAAATGCAATTTGATTCAGCAATCCCAATTA 2437

```

RESULT 8
HSN804900 2231 bp mRNA linear HTC 03-AUG-2004
LOCUS Homo sapiens mRNA; cDNA DKFZp686C0964 (from clone DKFZp686C0964).
DEFINITION AL833587
ACCESSION AL833587.1 GI:21734233
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM

REFERENCE
1 (bases 1 to 2231)
Bahr, A., Lauber, J., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M., and Wiemann, S.
The German cDNA Consortium
Direct Submission
Submitted (13-JUN-2004) MIPS, Ingolstaedter Landstr.1, D-85764

COMMENT
Neuberger, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp686C0964) is available at the RZPD Deutsches
Ressourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
<http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp686C0964>
Further information about the clone and the sequencing project is
available at <http://mips.gsf.de/projects/cdna/>.

FEATURES

SOURCE

1..2231
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp686C0964"
/db_xref="taxon:9606"
/clone="DKFZp686C0964"
/tissue_type="endometrium"
/clone_lib="686 (synonym: hlec3). Vector pSport1_Sfi; host
DH10B; sites SfiI + SfiI"
/dev stage="adult"
/note="genomic"

ORIGIN

Query Match 8.9%; Score 769.2; DB 3; Length 2231;
Best Local Similarity 73.5%; Pred. No. 2.6e-160;
Matches 1241; Conservative 0; Mismatches 263; Indels 184; Gaps 13;

CDS

```

/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
1..6966
/notes="Liver regeneration-related"
/codon_start=1
/product="LRG00057"
/protein_id="AA096270.1"
/db_xref="GI:3765430"
/translation="MGLCVLVASIAACPWGHSPSPVVDITTKGKLVSLGSEFT
QVAVFLGVPPAKPLGLFAPPEPABEWSFVKNTTTPVTPGESHGNAHLDOL
AALRWQNDINAFNGNDPVITIFGESAGSVSVLMIALSCNNTSFAAMQCLQK
TEALBELTVELKNDMSVTDVGLVLPKPEELTEKSNTPYIYKPGWGLMIP
TNISSEVPAIIETKLRGTHDPKAKKELLDMFSDVPRFPAVMSRSIRDGAADPM
YEFQYRPSFVDHPQTVQGDHDEI FSYFTGTPFLKEDGFDOFSVPSSADIDSR
FLQTLAVFLCEALCSFLQGECEQDSEKQKQSKVGLLECDI FTPRKGAMDE
HGQORSKLMRLKYKASGLPWGAORODCLHPRLKEDMINSILHSCRCRCEPLGST
SERQTSINQKRLPAHTRSPBEKAQDLBPMCTEAPRGGTGLPGCCRCRCEPLGST
PRANLSITTEIAEOSNTRDMARGKRPNRNQDMSSESPNSPTKTNMEYPTP
EKQDVSYSYLMLEDPKQMTLRETOENLNKQVAYRBSOKSLKEFOENTIKOL
KELKEITAIKEHEMTTIDINQKQGVADTSFTNRLQEMBERISGABDEIIDS
TKEDNVKRRKLVONI QETQDSMRSNRL IGIESBDSPLGPNVIFNKIIEENP
LKKEIPIDIOEAYRTPNRIDQKNTSRHIVTPNAONKERILIKAREGOVYKGRP
IRITPDPSEPTWKARSWTDVITLREHNCQPLLYPALISINIDETKIFDKTFT
OYSTNPALORINERFENISKOVEAHRETOKSLKEFOEKIKOLKELMEIBAIK
EHMETTIDINQKQGVADTSFTNRLQEMBERISGABDEIIDSIVYDNRKRLV
VONIDSKRSNRLIGIESBDSPLGPNVIFNKIIEENPFLKKEITGLOEAYRT
PNRLDQKNTSRHIVTPNAONKERILIKAREGOVYKGRKTPFQANQK
QAVSILISNKNIPOLKVKDKREHPIFKGKHODELSINIYAPNRAPIYVET
LAKTILAPHTIIVGDEPTPLSMDRSWKQILNSBVDLREVMQMDLTDIYRTP
KAKGYTPSAPHPGFEKIDHIIIGQKGLRKYRIEIPCVLDHSGLKVFNKGM
PYTWMKNNALINDLVKEIKELIKNLEPENEDTYSNIMDTKALFNNKGLMS
ACRKKOEAYVSLTAHLKALHOKENATPRSRROEIKLRAEINOVETKRTTERLN
TKMEFEKINKIDKPLARLGRHREVQINKIRNEGDIITVSEEOKIRSYKNLY
STKLENLOEMDFLDRYOVSKNOBQINDNPIPKELIYVAGLPKTSKSGPDS
AEFYOTFIEDLPILSKLFKIEEDGALNSFEYEAITIIIPKHOTKKNRPSLM
NINIDAKILIANRIOEHKTIILHDDQVPIPGMGWINKITIMHIVIKKLEON
HMITSLDAKAPDKOHPMIVLVERIGQBYINIVKAYSEKPAVNIILNKEIEAI
PLKSTROGCSPLPLNIVLVARATVYKQEKIKQIGKEBEKISLPAHMDIYVLS
DPKSTSELKLIINFSKVAGYKINSNSVAFLYKKEIEIREITRPIIDPNNIK
YLVGLTKQVODLVKNKFTLKEIEEDRWMDLPCSGVIGRINIYKVALIPRAHIF
NALPIKIDIOFEKELDRITCKEILANNKERRAKAILNNRSGGIIITPLKQYVRAIF
IKTAAVYWRNQIIDOMNRIEDENMHTYGLIIPDKGATIMOKOSIPSKQPMWR
ATCRMOIDPCSPTCKLKSXKIKDLHITPDLKILEEGHLMHEMGKQKPNKTP
MAYALRSRIDKDLIKLQSFCAKQTVVTRKQPTDMEKIFNPTTERELISKITKEL
KLDLRETNNPILKMGSELNKEFTLAEBCMAERHLKCGSTSLIIRQMVIYKTEL
PCDMLRSKTOVADAGEDEVEKESHSIVGRNWKERPCSTIEIRIOMVIYMEYVSA
IKNNEFMKFGKMLENIIISLELTOSODIHAPLYETLIRKGHFLYCOGQGVVKA
RAPOSILNSAREELPHFIKQFVLRFHGIIQNQNHSLGLSSSWPARSGSISTQDRA
FLHSRGEAVGAEVW"

```

ORIGIN

Query Match 8.7%; Score 749.2; DB 3; Length 6966;
 Best Local Similarity 65.0%; Pred. No. 9.5e-156;
 Matches 1274; Conservative 0; Mismatches 653; Indels 33; Gaps 10;

```

QY 1 TGTATGAAGCCAAATGCTACTTTAATACAAAACGAGAAAGATATA--CAAAAAGAAA 58
DB 4577 TCTACGAAGCCCAATTACTCTTATCTTAACCAACAAAGACACAAAGAAAAGAG 4636
QY 59 ACTATAGACCACTACGATGATATATACATGACAGAAATCCCAACAAAATCTAGCTA 118
DB 4637 ACTTCAGACCAATTTCCCTTAAGAAATATGAGCGAAAATAATCAATTAATAATCTGCA 4696
QY 119 ACCCAATCCACAGATATCAAGAGATATCCACCAATGTCAGTGGGTTTCATACCG 178
DB 4697 ACCGAATTCAGAGCAATCAAAAACAATCATCCACATGATCAAGAGGTTTATCCCG 4756
QY 179 GAGTGCACAGATAGTTAAACATACACAGTCAATAATGTGATATACATCAATAAGAA 237
DB 4757 GCATCAGGAGATGTATTATATACGAAAAACCATGATCATCTTAATATTAAGAAAC 4816
QY 238 TTAATAAACAATAATCATGATCATCTCAATAGATGTAATAAGCAATTTGACAAAATCT 297

```

```

DB 4817 TGAAGAAACAGAACCAATGATCATTTTATGATGCTGAGAAAGCATTTGACAAAATTC 4876
QY 298 AACATTTCTTTATGATTAATAACCTTCACGAAATGACATGAGAAAGCATTCCTTAATG 357
DB 4877 AACACCCCTTCATGATTAATAAGCTCGGAAAGAAATGGAATTCAGGCCCATCTTAACA 4936
QY 358 TATATAAGCCATATATATGACGAGCCACAGCAACATTAATATGATGAGGAGAAAGTTGA 417
DB 4937 TGTATAAGCCATATATGACGAAACAGATGCTTACATTAATTAATGAGAGAACTTG 4996
QY 418 AACATTTGCTCGGAAAGCTGGAACAAGCAAGATG--CTACTTTCACCACTTATTC 476
DB 4997 AAGCATCCCATTAATAATGAGGACGTAAGCAAGGCTGCGCACTCTCCCTCACTATTC 5056
QY 477 ACATAGTAGTGAAGTTTATGACGACATATGACAAAGAGAAAGAAATCAAGGACACC 536
DB 5057 ATATAGTTCTTAAGTTCTAGCCAGAGCATATGACAAACAAAGAGATCAAGGATAC 5116
QY 537 AATCAATAAGAGAAAGTCAACCTGCTGTCATGATGATGATGATGATGATGATGATGATG 596
DB 5117 AGATGGAAGAAAGAGAGTCAAAATATCACTATTGACATGACATGATGATGATGATGATG 5176
QY 597 AAAACCTTAAGACTCATCTACAGAAAGCTCTAGAACTGATACATAATTCAGTAAGTTT 656
DB 5177 GTGATCCCAAAAGTTCCACAGATGAACTATGAAGTATGATGATGATGATGATGATGATG 5236
QY 657 CAGGATTAACAATTAATGTAACAATACTGACGCTGTATGACCAAGTACCAAG 716
DB 5237 CTGGGATATTAATTAATCACTAAATTAATCAGTTGCTCTCTACCAAAAGAGAAACAG 5296
QY 717 CTGAGAACTAATTAAGAACTCAACACCTTACATAGCTGTAA-----AAATCT 770
DB 5297 CCGAGAGAAATTAAGGAAAGACACCTTATATATGACCAATTAATATTAATTAATTAAT 5356
QY 771 TAAGATATCTTACCCAGAGAGGTGAAGACCTTACAGAGAAATCTACAAACACAG- 829
DB 5357 TCGGTGTACTTTAACCAAGAGTAAGTAAGATCTGTGATGATGATGATGATGATGATGATG 5416
QY 830 ---CTGACATCATGATGACACAAACAGTGAACACATCCCATGCTCATGATGAGTGA 886
DB 5417 GGAAGAAATTAAGAAAGCTCAGAAAGATGAAAGATCTCCCATGCTCATGATGAGTGA 5476
QY 887 GATCAATCTGTGAAAGAAAGACATATGTCGAAAGCAATCAAGATTCATGCAATTC 946
DB 5477 GGAATTAATTAAGTAAGAAAGTGCATTTTACAAAGCAATCAAGATTCATGCAATTC 5536
QY 947 CCACCAATATCATCATCATCTTTCACAGAACTGAGAAAAACAATTCATAATTCATA 1006
DB 5537 CCATCAAAATTCATCAATCAATCTTTCAGAAAGATTAG--ACAGAACAATTTGCAAAATTC 5595
QY 1007 TGGACCAACACCAAAAAAAAAAAAAAAAACCCGATAGCCAAAGCAAGACTTAGCAAAA 1066
DB 5596 TGGAAATAC-----AAAAACCAGATAGCTAAAGCTATCTCCCACAATA 5641
QY 1067 AGAACCAATCTGAGGATCATATTAACCATCTTCAAACTATACATCAAGGTTATATCA 1126
DB 5642 AAAGACTTCAGAGGGAATCACTATCTCGAAGCTCAAGAGATTAACAGAGAAATAGGA 5701
QY 1127 CCAAAATCATGAGCATGATTAATAAATAGGACATAGACCAATGAAAGAAAGAGAGA 1186
DB 5702 TAAAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 5761
QY 1187 ATCCAGAAATTAAGCCAAATTAATTAAGCAACTGATTTTGAACAAGCAACAAAACA 1246
DB 5762 ACCCAAAATTAAGCAACCACTATGATGATGATGATGATGATGATGATGATGATGATG 5821
QY 1247 TAAAGTGGGAAAGACATCTTATGATTAAGAAATGCTGAGATTAATGCAAGCAAT 1306
DB 5822 TCAATGAGAAAGAAAGATGACATTTTCAGCAAAATGCTGATGATGATGATGATGATGATG 5881
QY 1307 GTGAGAGATGAAAGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1366

```

Db 5882 GTAGAAGATGACATGATTCATGCTC-ATCACCCCTGTACAAAGCTTAAGTCCAGG 5940

Qy 1367 ATCAAAGACTTAAATCTGAGACCTAAAACATTAATAATTTAGAGATATACATGAGAAA 1426

Db 5941 ATCAAGACCTCCACATCAACAGACACACTCAACTTAATAGAGAAAAGCTAGGAG 6000

Qy 1427 ATGCTTACAGATTCATTCAGGCAAGACTTCAATGGCCAAAGAACCCAAAGTAAATGCA 1486

Db 6001 CATCTGAGACACAGGGGCACTGGAATAAATTTCTGAAACAAACCAATGCTTATAGCT 6060

Qy 1487 ACAAAAACAAAATAATATAGATGAGACTTAATTAATTAACCTTAATTTGGCAGCAAAA 1546

Db 6061 CTAAGATCAAGATTCGACAAATGGGATCTCAATAAAGCTGCAAGCTTCTGTAGGCAAG 6120

Qy 1547 ACAATCTTAGCAGACGACAAACCAACCCAGAGAGAGAAAATCTTCAACAACTTA 1606

Db 6121 GACACGGTGTAGGACAAACGAGCAACCAAGATGGAGAAAAGATCTTTACCAATCCT 6180

Qy 1607 GCATCTACTAGACATTAATTCGGAAATCCAGAAAGAACTCAACAAATACGACAAAG 1666

Db 6181 ACAACAGAGAGAGGCTTATATCCAAATATATCAAAAGAACTCAAGAGTTAGACCGAG 6240

Qy 1667 AAGCAACAATCCCATGAAAGAGTGGCTAAGACATGAATGACAAATTTCTCAAAAGAA 1726

Db 6241 GAACAAATTAACCTTATTAATAAATGGGTTCAAGCTTAACAAAGATTCACAGCTGAG 6300

Qy 1727 GATATACAAATGGCCAAACAAAC--GGAATAATGCTTAACATCAATATGATTAGGAA 1784

Db 6301 GAATGCCGAATGGCTGAGAAACACCTAAAGAAATGTTCAATCTTATAGTAAAGGAA 6360

Qy 1785 ATGTAATTCACATCTGTAATGCGATACCACTTACTCTGCAAGAAATGCTATATTTAA 1844

Db 6361 ATGCAAAATCAAAACCAACCTGAGATTTCACTCAACCC-ATGTAATGGCTTAACATCAA 6419

Qy 1845 AAATCTAAATAATATATGATGTTGGTGGCTGTGAGTAAAGAAACATTTACAGTGC 1904

Db 6420 AACTCAGTACAGCAATGCTGGCGAGATGTGAGAAAGAAAGAAACCTCTCCATTTGT 6479

Qy 1905 TGGTGGAAATGTAACTTGGCGCAACCACTATGAGAAAACAG 1944

Db 6480 TGGCAGAACTGGAAGAACCAACCATGATGCCCTTCAACAGAG 6519

RESULT 10
AY539894 3429 bp mRNA linear HTC 21-MAR-2004

LOCUS Rattus norvegicus LRRGT00143 mRNA, complete cds.

DEFINITION AY539894

ACCESSION AY539894.1 GI:45478125

VERSION

KEYWORDS HTC.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 3429)

AUTHORS Xu,C.S., Zhang,L., Chang,C.F., Han,H.P., Wang,G.P., Chai,L.Q.,
Yuan,J.Y., Yang,K.D., Zhao,L.F., Ma,H., Wang,L., Wang,S.F.,
Xing,X.K., Shen,G.M., Shi,J.B., Rahman,S., Wang,Q.N. and Zhang,J.B.

TITLE Liver regeneration after PH

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 3429)

AUTHORS Xu,C.S., Zhang,L., Chang,C.F., Han,H.P., Wang,G.P., Chai,L.Q.,
Yuan,J.Y., Yang,K.D., Zhao,L.F., Ma,H., Wang,L., Wang,S.F.,
Xing,X.K., Shen,G.M., Shi,J.B., Rahman,S., Wang,Q.N. and Zhang,J.B.

TITLE Submitted (03-FEB-2004) Henan Bioengineering Key Lab, Henan Normal
University, NO.148 Jianshe Road, Xinxing City, Henan 453002,
P.R.China

FEATURES

source location/Qualifiers

1..3429 /organism="Rattus norvegicus"

/mol_type="mRNA"

CDs

/db_xref="taxon:10116"

1..3429

/note="Liver regeneration-related"

/codon_start=1

/product="LRRGT00143"

/protein_id="AA566234.1"

/db_xref="GI:45478126"

/translation="MVOQTRSEVVISQWESODERKNYTFNPLLIQTSPSPNTVKLLN
SVDTGMEKHKEKVVNDIYIHSRCIFKPMKDKDSQVPLDISGSGIGDPPASGF
RAYVCARDITRONSFRDQCILGSLYXMRQYKLNQEQINQJNBI1PPEKEKAEKI
GLPTKSPGPDGSAEYOFIEDLAPILSKLPHKEITDGLPNSFEATYTL1PKPH
KDTTKENPRPISLMTIDAKILKIANLOEHIKTIYHNDQGFITGMOKPRT
IWIYHTINKLKEQNHMTISDAEKAPKTOHPMRUKLEIGTIGSPILNKAISRP
VANIKLNGEKLEAIPLKSGTRGOCPSLPYFNIVLELAARIRQEKIKGIQIEKEV
KISLFPADMTIVLSDPSSTRELLKILNNSKVAQYKINSKSYALPYTEKQAEKI
RETPPFIIDNNIKYLGVTLLKQVLDYNNKFLREIEEDLARMOLPCSMGRIN
IVMAALPKAIYFPNAPIKIPLOFELRLTCKEFLMNKKPPIAAILNNKTSGG
IT1PELKQYRAIVITKAWYVRDROIDONR1EDPEMNPHTYGH1EDGAKTIOWK
KDISPSKCMFMRSTCRMQIDPCSPCTKLSKTKID1H1KPTDKLIEKRGKL
EHNGTGNPLNKTPMAYALRSRIDKMDLILQSFCKADVTVTRKOPTMEK1FTNP
TDRGLISKLYKELKLDRETNPNPIKMGSELNKOPTAECEVAEHLKMKCS1LV1
REMOIKTTLFPHLTPVMMAKIKNSALF1IARKKEPCSPTEEWI1QOMWY1YTWBY5
AIKNNSMKFVGMLELENI1LSELNOSORDIGGHLI1IARVSHGLMTVGNVLTG
GWSMTAPILGDHSGDGECELOVNGAOGYODVTP1OGAKTLHP1PMPPIKPKAK
RIENHSRRATKQOBBSVCKMEDLLVNAITLKTFTILIRKGRYILASCIARIGINT
FFFFFVRLHPONPCENAGDGPVAVLACSMQISRYRPSSESPVSNDQSKNA
OAAENRRETFMGDNDVAVDGSSTENKSNRQELH1QIRASGHORVTSWKAT"

ORIGIN

Query Match 8.7%; Score 747.8; DB 3; Length 3429;

Best Local Similarity 66.0%; Pred. No. 1.7e-155;

Matches 1231; Conservative 0; Mismatches 602; Indels 32; Gaps 9;

Qy 1 TGTATGAAGCAATGTCACTTAAATACCAAAACAGAGAAAGATATA--CAAAAAGAA 58

Db 617 TCTGAGAGCCACATTAATCTTATACCTTAACCAACCAAGACACAAAGAAAGAA 676

Qy 59 ACTATGACCGATACCACTGATGAATATATGACGAATATCCCAACAAATTAAGCTTA 118

Db 677 ACTTACAGACCAATTTCCCTTATGAAATATGACGCAAAATAATCTCAATTAATTTGCA 736

Qy 119 ACCCAATCCACGACATATGAGAAATATCCACCTTGTCAAGTGGTTTCAATACAG 178

Db 737 ACCGAATTCAGAGACATCAAAACATATCCACATGATCAAGTATGCTTCAATCCAG 796

Qy 179 GGGTGA-GGATAGGTTAATACATACACAAATGAATTAATGATATCATCATTAACAGAA 237

Db 797 GCATGAGGAGATGTTTAATATGAGAAACCATCATATGATGATCATTAATTAACAAAC 856

Qy 238 TTAATAACAAAATACATGATATCTCAATGATGCTGAAAAAGCAATTTGACAAATCT 297

Db 857 TGAAGAAGCAAGAACCATATGATATTTCAATTAAGTGTGAAAAAGCAATTTGACAAATTC 916

Qy 298 AACTTTCTTATGATTAATTAACCTTACGAAATTCACATGAAAGACATACCTTATG 357

Db 917 AACACCCCTTCATGATTAAGAGCTCTGAAAGAAATGGAATTCAGAGCCCATATCCTTAACA 976

Qy 358 TAAATAAGCATATATATGAGAGACCCACAGCAAAATTAATATGATGAGGAAAAAGTTGA 417

Db 977 TAGTAAGGCAATATACAGCAACCAAGTGTCTAATTAATTAATGAGAGAAACCTCG 1036

Qy 418 AAACATTTGCTCTGAGAACTGGAACAGCAAGAGAG-CTACTTTCAACACTTATTTCA 476

Db 1037 AAGCAATCCACATTAAGTACAGAGGCTGAGCAAGGCTGCCACATCTCTCCCTACTTAATTTCA 1096

Qy 477 ACATAGTACGAGAGTTTACCGAGAGCATATGAGCAAGAGAAAGAAATCAAGGCAACC 536

Db 1097 ATATAGTCTTGAAGTCTGAGCAAGCAATGAGCAACAAAGAGAGATCAAGGAGATAC 1156

Qy 537 AAATCAATAAGAGAGAGTCAAACTGCTCTGTTCACTGATGATATGATATGATACCTAG 596

Db 1157 AGATCGAAAAGAGAGAGTCAAAATATCACTATTTGACAGATGACATATGATATTTAA 1216

QY 457 ACTTCACCACTTCTATTCAATAGTAGGGAAGTTTACGACGAAATCAGACAGA 516
 Db 2742 ACCCTCCACCACTGTACTCAATTTAGACTGGAAAGCTCTAGCTGAGTAATCAACAGAGA 2683
 QY 517 GAAAGAATCAAGGCAACCAATCAATTAAGAGGAAGTCAACCTGCTCCCTGTTCACTGA 576
 Db 2682 GAAAGAAATTAAGGCAATCCCAACCGAAAGGAAGACATCAAAAGTATCTGTTGTGAGA 2623
 QY 577 TGATATGATGTATATCTAGAGAAACCTTAAGACTCATCCGAAAGCTCTAGAACTGAT 636
 Db 2622 TGATATGAATTTATATTTGGAAAAACCTTAAGACT--ACCAAAACCTGTCAGTAGTAT 2565
 QY 637 ACATTAATTCAGTAAAGTTTACAGATACAACTAAATGTACACAATATCAGTAGACATGCT 696
 Db 2564 AAACAAAGTTCAATAGTTGCAAGATTAATAATCAACAAAAATTCAGTAGATTTCT 2505
 QY 697 ATACACCAACAGTAGACCAAGTGAAGATCAATCAAGAACTCAACACTTTTACAAATAGC 756
 Db 2504 ATATGCCAAGGTGAACATCTGAAAAACAAATTTAAATTAACCCATTTATATAGTGC 2445
 QY 757 TGTAA--AAAAATCTTAAGATATTTCTTACCAAGAGGTGAAGAGCTCTACAGAGAA 814
 Db 2444 CACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2385
 QY 815 AACTACAAACACAGCTGACA-----TCATGATGACACAAACAGTGGAAACATCTCC 869
 Db 2384 AAATATTAACATGATGAAGAAATGAGAGACACAAAAAACTGAGAGATATTTCC 2325
 QY 870 ATGCTCATGATGGTGAATCAATATTTGAAAAATGACATATTTGCCAAAGAACTGA 929
 Db 2324 ATGTTCATGATGAAGAAATCAATATTTGAAAAATGCTCACTAGCAACCAATCTA 2265
 QY 930 CAAGTTCATGATTCCTCCACCAAAATATCATCATTTCTTCCAGACAGTAAAAAAA 989
 Db 2264 CAGATTCATGATTCCTCTATCAAAATACAGTACATTTCTTCCAGAGATAG--GAAAA 2206
 QY 990 CAATTCATTAATTCATATGGAACACACAAAAAATTAATTAATTAATTAATTAATTA 1049
 Db 2205 GAATCCAAAAATTTATATGACGAC-----AAGAGACCCAGAAACAGCCAA 2160
 QY 1050 AGCAAGCTTAGCAAAAAAGAACAAATCTGAGGATCACTATCCCATTTCAAACTATA 1109
 Db 2159 AGCTATCTTAACAAAAAGAACAAACCTGAAGAACCACTACCTGCTTCAATATATA 2100
 QY 1110 CTACAGGCTATATCAACAAACATCATGCTGACATATAAATAGGACATAGACCA 1169
 Db 2099 CTACAGGCTATATTAATGAAATGACATTTGTACTGGCATTAATAACAGACATAGACCA 2040
 QY 1170 ATGAAAAAGAGAGAAATCCAGAAATTAAGCCAAATTAATTAATTAATTAATTTTGA 1229
 Db 2039 ATGGAACAGAAATAGAGAACCCAGAAACAAATCCACATGACAGTGACTCATTTTCAG 1980
 QY 1230 CAAAGCAACAAAAACATTAAGTGGGAAAAAGACATTTCTAGTTAAACAAATGTGCTGAGA 1289
 Db 1979 CAAAGGTGCAAGAAACATATCTGCGGAAAAAGACAGTCTTCAATTAATTAATGTCCAGGGA 1920
 QY 1290 TTATTTGCAAGCCATGTGGAAGATGAAGCTGATCCCTGTCTCTCACTTAATTAACAA 1349
 Db 1919 AAATCTGATATCTATATGCAAGAAATGAATCTAGA--CCCTATATCTTCCACATATACAA 1861
 QY 1350 AAATTTATCAAGATGATCAAGAACTTAATCTGAGACCTTAACCAATTAATTAATTTCTAG 1409
 Db 1860 AAATCAATCAAG--TAGATTAAGACTTAATCTGAATCTCAAACTATAGAACTATGAT 1802
 QY 1410 AAGTAACATCAAGAAAAATGCTTTTACGATTCCTTAGGCAAAAGACTTCAATGGCCAGA 1469
 Db 1801 AAGAAAAATTTGGAAAAATCTCCAGGACATTTGCTGGGCAAAAAATTTTGTGATATA 1742
 QY 1470 ACCCAAAAGTAAATGCAACAAAAAATTAATTAATTAATTAATTAATTAATTAATTA 1529
 Db 1741 CCCACAA--TCACAAACAAACAAAGGCAAAATGACAAATGGGATCAATTAATTAATAA 1683

QY 1530 GCTTTGGGAGCAAAAAACATCATTTAGCAGACCAAGCAACCCACGAGTAGAGA 1589
 Db 1682 CTTTCTCAGCAGCAAGAAAAACATCAACAAATGAAATTAACAAACCCGACGAATGGAGA 1623
 QY 1590 AAATCTTCAACAACTTAAGCATCTGACTAGGACTTAATATCCGAATCCACAGAACTCA 1649
 Db 1622 AAATATCTGCAACTACCCATTTTGACAAAGGTTAATTAATTAATTAATTAATTAATTA 1563
 QY 1650 AACCAATCAGCA--AGAAAGAAAGCAACATATCCATGAAAGATGGCTTAAGACATGAAT 1708
 Db 1562 AACCAACCTGTAGAAAAAATCTTAATTAATCAATCAAAAAATGGGCAAAAGATTTGAAAC 1503
 QY 1709 AGCAATTCCTCAAAAGAGATATCAAAATGGCCAA--AACAGAAAAAATGCTTAACA 1766
 Db 1502 AGATATTTCTCAA--AAGACATACAAAGTGGCAACAGCATTAAGAAAAGGTGCTCAATA 1446
 QY 1767 TCATTAATGATTTAGGAGAAATGTAATCAACACTGTAATGCGATACCACTTACTCTGCA 1826
 Db 1445 TCATGATCATCAGAAAAATGCAAAATGCAAACTGAATGAGACATCATGACCCAGTT 1386
 QY 1827 AGAATGCTCATTAATTTAAAAATCTAAAAATTAATAGATGTTGGGTCTGTGTGATPAA 1886
 Db 1385 AAATGCTTACATCCAAATATACAACTAATTA--TGCTGTGAGATGTGTGAAAG 1330
 QY 1887 GGAACACTTTTACATCTGCTGTGGGAAATGTTAACTTGCGCAACCATATGAAAAACGTG 1946
 Db 1329 GTATCTTCTGTACATCTGTGTGGGAAATGTTAAATTAATTAATTAATTAATTAATTA 1270
 QY 1947 TGGAAATTTCT--TAAGAACCTAAAGATGATGACATTTGATCAGCAATCCCATTA 2003
 Db 1269 TGGAGTTCTCTCAAAAAAATTAATAATAGACTTCAATGATCAGCAATCTCACTA 1212

RESULT 12
 AK038045
 LOCUS
 DEFINITION
 Mus musculus 16 days neonate thymus cDNA, RIKEN full-length
 enriched library, clone: A130073006 product: unclassifiable, full
 insert sequence.
 AK038045
 AK038045.1 GI:26086252
 HNC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL
 MEDLINE
 PUBMED
 20499374
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Komno, H., Akiyama, J., Nishi, K., Kitanishi, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matshiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, Y., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL
 MEDLINE
 PUBMED
 20530913
 11076861


```

QY 1373 GACTTAATCTGAGACTTAAACCATTAATTTCTAGAGATAATCATCAGAAAATGCTT 1432
DB 1365 GAATCTCATATTAATAATGAGACAGTGAACCTTTAGGGGAGAAAGTGGGAGAAAACCTT 1424
QY 1433 CTAGACATTCATCATTAGGCAAGACTTCAATGGCCAGAACCCAAAAGTAAATGCAACAAA 1492
DB 1425 GAGATTTTGGCAGACGGGGAAAAATTCCTGTAATGAAACAGCAATGGCTTGCTGTAGA 1484
QY 1493 ACAAAATTAATAGATAGACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1552
DB 1485 TCGAGATTGTCAGAAATGAGACTCATGAACTGCAAGCTTCTGTAAGGCAAAAGACACC 1544
QY 1553 ATTGACGAGCAACAGACAAACCAACCGAGTGAAGAAATCTTCAAACTTAAGCATCT 1612
DB 1545 GTCAATAGACAAAAGGCCCAACAGATTGGGAAAGATCTTACTATCTTAAGTCA 1604
QY 1613 GACTAAGACTTAATATCCGAAATCCAGAGAACTCAACAAATCAGCAAGAGAAAGCA 1672
DB 1605 GATGAGAGACTTAATATCATATATATTAATTAAGAACTCAAGAAAGTGAATCCAGAAATCT 1664
QY 1673 ACAATCCCA-TGAAGAGTGGCTTAAGACATGATAGCAATTTCTCAAAAGAGATAT 1731
DB 1665 AATTAACCCATTAAATAATGGGCTCAGAGCTAAACAAAGATTTCTCACTGAGGATC 1724
QY 1732 ACAATGGCCCAACAAAC--GGAAAAATGCTTAACATCACTAATGATTAGGAAATGTA 1789
DB 1725 CCGAATGGCTGAGAAACACCTGAAAAAATGTTGAGCATTAATCAATCAGGAAATGCA 1784
QY 1790 AATCAACACTGTATATGGGATACCACTTACTCTGCAAGAAATGGCATTAATTAATTAATC 1849
DB 1785 AATCAAAACAACCTTGATGATTCATCTCAACGCTCAAGATGGCTTAAGATCAAAATTC 1844
QY 1850 TAAAAATTAATAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1909
DB 1845 AGGTGACAGAGATGCTGGGAGAGATGAGAAAGAGAAACATCTCTCCACTTTGTTGTTG 1904
QY 1910 GGAATGTAACCTTGGCAGCACTATATGSAACACGTTGCAAAATTTCTTAAGAACTTAA 1969
DB 1905 GGAATGTAACCTTGTATACCACTCTGAAATCACTGTCGCAATTTCTCAAGAAATTTGA 1964
QY 1970 AGTAGATGACCATTTGATCCAGCAATCCC 1999
DB 1965 CATATGTAACCGAGGATCCGCAATACC 1994

RESULT 13
CA455045 889 bp mRNA linear EST 12-NOV-2002
LOCUS AGENCOURT_10735784 MAFEL Homo sapiens CDNA IMAGE:6722551 5',
DEFINITION mRNA sequence.
ACCESSION CA455045
VERSION CA455045.1 GI:24905369
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 889)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@rs-remail.nih.gov
Tissue Procurement: Kristi A. Egland, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM14285 row: h column: 07
High quality sequence stop: 743.

```

```

FEATURES
... SOURCE
Location/Qualifiers
1. 889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722551"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231,
hTERT-HME1, LNCaP"
/lab_host="EMD10B"
/clone_id="MAPCL"
/notes="Vector: pCMV-Sport6; Site 1: EcoRV; Site 2: Not I;
Subtracted with brain, liver, lung, kidney and muscle.
Directionally cloned. Priming method: oligo-dT. Average
insert size: 1800 bp. Library amplification: 26,000 fold.
Kristi A. Egland, James J. Vincent, Robert Strausberg,
Bungkook Lee & Ira Pastan: Discovery of new breast
cancer genes encoding membrane and secreted proteins.
Manuscript submitted."

ORIGIN
Query Match 8.2%; Score 708.2; DB 6; Length 889;
Best Local Similarity 97.8%; Pred. No. 9.4e-147;
Matches 727; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 7502 GGGTCATCAGACTGGGGTTTCTGAGCATGATTAACCATCCAGTCTGGGTACAGAAC 7561
DB 147 GGGTCATCAGACTGGGGTTTCTGAGCATGATTAACCATCCAGTCTGGGTACAGAAC 206
QY 7562 TGACCAATCAACGAGAGTGAAGAGACTCTGCTCAAGCAGACCTGAGCTTCAAGG 7621
DB 207 TGACCAATCAACGAGAGTGAAGAGACTCTGCTCAAGCAGACCTGAGCTTCAAGG 266
QY 7622 GGGTCATCAGACTGGGGTTTCTGAGCATGATTAACCATCCAGTCTGGGTACAGAAC 7681
DB 267 GGGTCATCAGACTGGGGTTTCTGAGCATGATTAACCATCCAGTCTGGGTACAGAAC 326
QY 7682 TGAGGCTCCGATGCGGAGAGAGCTGTCTCATCTACATCTCAACCTGTCCGGCCG 7741
DB 327 TGAGGCTCCGATGCGGAGAGAGCTGTCTCATCTCAACCTGTCCGGCCG 386
QY 7742 ACTTCCTCTCTTAAAGGAGCACTATATGTTGCGGCTTAAGCCTCATCAATATCCGGC 7801
DB 387 ACTTCCTCTCTTAAAGGAGCACTATATGTTGCGGCTTAAGCCTCATCAATATCCGGC 446
QY 7802 ATCCCATCTCAAAATCTCAGTCTGTGATGACCTTTCCTTACTTATAGGCTTAAGCA 7861
DB 447 ATCCCATCTCAAAATCTCAGTCTGTGATGACCTTTCCTTACTTATAGGCTTAAGCA 506
QY 7862 TGCTGAGCGCATCAGACCGAGAGCTGCTGTCCATCTGTGCGCATCTGTGTAACA 7921
DB 507 TGCTGAGCGCATCAGACCGAGAGCTGCTGTCCATCTGTGCGCATCTGTGTAACA 566
QY 7922 GCGGCGCGCCAGATPACTGTCATCGGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 7981
DB 567 GCGGCGCGCCAGATPACTGTCATCGGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 626
QY 7982 TGCGAGATATCTGAGTGAATGTTCTGTGACTTCTGTTAGTGTGATTCGTTT 8041
DB 627 TGCGAGATATCTGAGTGAATGTTCTGTGACTTCTGTTAGTGTGATTCGTTT 686
QY 8042 GGTGTGAAGCTCAGATTTCAATCAATCCGCTGCTGTTTATATGTTGTTCTCT 8101
DB 687 GGTGTGAAGCTCAGATTTCAATCAATCCGCTGCTGTTTATATGTTGTTCTCT 746
QY 8102 GTGGGTGACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8161
DB 747 GTGGGTGACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 806
QY 8162 CAGAGCTGACGTGACCATCTCTCAACAGTGTGCTGCTCTCTCTCTGTGCTGCTGCTGCTGCT 8220
DB 807 CAGAGCTGACGTGACCATCTCTCAACAGTGTGCTGCTCTCTCTGTGCTGCTGCTGCTGCTGCT 866
QY 8221 TTGGCATTCAGTGGGCTCTGTT 8243

```


Db 867 TTTGGCATTTCAGGGGCGCTT 889

RESULT 14
LOCUS BC028293 2776 bp mRNA 1linear HTC 04-MAR-2003
DEFINITION Homo sapiens, clone IMAGE:4704802, mRNA.
ACCESSION BC028293
VERSION BC028293.1 GI:22418059
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2776)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (10-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
REMARK Contact: MGC help desk
COMMENT Email: cgabbs-remail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louissegh, H.,
Kovats, C.R., Sneed, A.J., Martin, R.G., Wuzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LIML at: <http://image.jnl.gov>
Series: IRK Plate: 29 Row: 1 Column: 19
This clone has the following problem: retained intron.
Location/Qualifiers
1. 2776
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4704802"
/issue_type="Breast", mammary adenocarcinoma."
/clone_lib="NIH MGC_87"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 8.2%; Score 702.8; DB 3; Length 2776;
Best Local Similarity 72.9%; Pred. No. 1.9e-145;
Matches 1009; Conservative 0; Mismatches 347; Indels 28; Gaps 7;

QY 628 AGAAGTGTATCAATTAAGTTTCAAGATCAAACTAAATGTACAAATCAGT 687
DB 864 AGAGTAGCTTCAAGCTTCAGCGAAGTCTCAGATATTAATCAATGTGCAAAAAATCACA 923
QY 688 AGCAGCTATACACCAAGAGTGAAGCTGAAGTCAATCAAGAACT--CAAGACT 745
DB 924 AGCATCTTATACACCAATTAACAGAGAGAGCCAAATCTAGTGAATCCCAATTACA 983
QY 746 TTACATATAGCTGTAAATAATCTTAAGATATTTCTTACCAAGAGGTGAAGACTTC 805
DB 984 ATTGCTTCAGAGATTAATATCTAGATTCATCAAGGATGGAGAGACTTC 1043
QY 806 TACAAAGAAATCTACAAACACAGCT---GACATCATAGATGACAAACAAAGTGGAAA 861
DB 1044 TTCAAGAGAACTACAAACACAGCTCTCAAGGAATTAAGAGATACAAACAAATGGAAG 1103

QY 862 CACATCCAGCTCATGATGGGTAGATCAATATTTGTGAATAATGACCATATGGCAGAAA 921
DB 1104 AACATTCATGCTCATGGGTAGAGAAATTCATTTCTGAAAATACCATATGCCCCAG 1163
QY 922 GCAATCTACAAATGTCATATGCCAACCAAAATATCATCATCTTCTTACAGAACTA 981
DB 1164 GTAAATTCAGATCAATGCCATCCCATCAAGTACCAATGACTTTCTTACAGAAA--TT 1222
QY 982 GAAAAAACAATTTCTAAATTCATATGGAACACAAACCAAAAAAATTTTTTTTTTCCCGC 1041
DB 1223 GAAAAAACAATTTCTAAATTCATATGGAAC-----CAAAAAAGCCAC 1268
QY 1042 ATAGCCAAAGCAAGCTTACCAAAAAAGAACAAATCTGAGAGGATCATACCTTCCATCTTC 1101
DB 1269 ATTGCCAATCAATCTTAAGCCAAAAGACAAAGCTGAGAGATCAGCTTCTGACTTC 1328
QY 1102 AAATCTACTACAAAGCTTATATCAACCAAAATCATGAGCTGACATPAAACTTGGAC 1161
DB 1329 AAATCTACTACAAAGCTTATATCAACCAAAATCATGAGCTGAGATCAGCTTCTGACTTC 1388
QY 1162 ATAGACCAATGGAAGAAAGAGATCCAGAAATTAAGCCAAATTAATTAAGCCAACTG 1221
DB 1389 ATAGACCAATGGAAGAAAGAGATCCAGAAATTAAGCCAAATTAATTAAGCCAACTG 1448
QY 1222 ATTTTGAACAAGCAAAACAAACATTAAGTGGGAAAGACATTTCTAGTTAACAAATG 1281
DB 1449 ATCTTTGACAAACCTGACAAACAAACAAAGATGGGAAAGATTCCTTATTTTAATTAATG 1508
QY 1282 TGTGATATTTATGGAAGCCACATGTGGAAGATGAATCTGATTCCTTCTCTCACT 1341
DB 1509 TGTGGAAGAACTGCTGATGACATATGTAGAAAGCTGAATCTGATTCCTT--CTTTACACC 1567
QY 1342 TAATCAAAATTTGATCAAGATGATCAAAACCTTAATCTGAGACCTTAACCTAAACATTA 1401
DB 1568 TTAATCAAAATTTGATCAAGATGATCAAAACCTTAATCTGAGACCTTAACCTTAACATTA 1627
QY 1402 AATTCTAGAAATTAATCATGAGAAATATGCTTTAGACATTCCTTAGGCAAAAGCTTCAT 1461
DB 1628 AGCCCTAGAAAGAAACCTTAGGCAATACATTCAGACATAGGCAATAGGCAAGAGCTTCAT 1687
QY 1462 GGCAGAAACCCAAAGTAAATGCAACAAACAAATTAATAGATGAGCTTAATTTA 1521
DB 1688 GTCTAAACACCAAAAGCAATGAGCAAAAGCCAAATTTGCAATATGGATCTTAATTA 1747
QY 1522 ACTAAACCTTTTGGCGAGCAAAACAAATCTTGGAGAGAAACAGAACCCAGCA 1581
DB 1748 ACTCAAGACCTTCTGACAGCAAAAGAAAGCAACCTCAGAGTGAACAGCAACCTTACGA 1807
QY 1582 GTGAGAGAAATCTTCACAAACTTAAGCATCTGACTTAAGGACTTAATATCCGAAATCAGAA 1641
DB 1808 ATGGAGAAATTTTGGCAATCTACTCANTCTGACAAAGGCTTAATATCAGAAATCTACA 1867
QY 1642 GGAATCAAAACAAATCAGAGAGAGAAAGAAACAAACATCCATGAAGAGTGGGCTTAAGA 1701
DB 1868 TGAATCAAAACAAATTTACAGAAAA---AAACAACCCATCAAAAGGTGGCAAGAA 1923
QY 1702 CATTGAATAGCAATTTCTCAAAAGAGATTAACAATGGCCAGCA--ACAGGAAAAAATG 1759
DB 1924 TATGAACAGACCTTCTCAAAAGAGACCTTTATGCAAGCAAAAAACATGATAAAAAATG 1983
QY 1760 CTTAATCATCATATGATTAAGGAAATGTAATTAACAACCTGTAATGCAATACCACTTAC 1819
DB 1984 CTCTATCATCTGCGCATAGAGAAATGCAAAATCAAAACCAATGAGATACCATCTCAC 2043
QY 1820 TCCTGCAAGATGCTATTAATTAATAATCTAATAATTAATAGATGTTGGTGGTCTGTGG 1879
DB 2044 ACCAGTTAAGATGCAATCATTAATAAAGTCAGAGAAACAAACAGGTCTGGAGAGATGTAG 2103
QY 1880 TGAATAAGGAACATTTTACATGCTGTGGGAAATGTAACTTGGGCAACCATTAATGAA 1939
DB 2104 AGAAATAGGAACATTTTACATGTTGGGGAATGTAACTTAAGTCAACATTTGTGAC 2163
QY 1940 AACAGTGTGAATTTCTTAAGAACTAATAAGTATGACATTTGATCCAGCAATCCC 1999


```

Db      2164  GTCACTGCGGATTCCTCAGAGATCTTGAACCTAGAAATACCATTTGACCCAGTCATCCC 2223
Qy      2000  ATTA 2003
Db      2224  ATTA 2227

RESULT 15
CN839125
LOCUS
DEFINITION
AGNCOURT 15864180 NIH MGC 145 Homo sapiens cDNA clone
IMAGE:30706609 3', mRNA sequence.
ACCESSION
CN839125
VERSION
CN839125.1
KEYWORDS
GI:47944780
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 871)
AUTHORS
NIH-MGC http://mgi.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgaabds-romail.nih.gov
Tissue Procurement: GPCR Consortium
CDNA Library Preparation: GPCR Consortium
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
plate: IRB14 row: e column: 02
High quality sequence stop: 701.
Location/Qualifiers
1..871
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30706609"
/tissue_type="mixed"
/lab_host="DH10B"
/clone_lib="NIH MGC 145"
/notes="Vector: pCDNA3.1; Site 1: varies by clone; Site 2:
varies by clone; ORFs were PCR-amplified and cloned into
pCDNA3.1 by the GPCR Consortium. Cloning sites vary by
clone and include the following: 5'-EcoRV-XmnI/XhoI-3',
5'-EcoRV-XmnI/NotI-3', EcoRV (TA cloned, non-directional).
For information about which gene each clone represents,
please visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat
a Note: this is a NIH_MGC Library."
ORIGIN
Query Match 7.9%; Score 683.4; DB 7; Length 871;
Best Local Similarity 89.6%; Pred. No. 3.2e-141;
Matches 757; Conservative 0; Mismatches 86; Indels 2; Gaps 2;
Qy 7309 AGTCAGGCTTCAGAGTCACAGAACTGATTTCAACTGATTTGAGACCCCACT 7368
Db 7 AGTCAGGCTTCAGAGTCACAGAACTGATTTCAACTGATTTGAGACCCCACT 66
Qy 7369 TTGGATAGGTCATTTCTCTGCGAGTCTTGATCTCTCTTTAAATAGACAGTA 7428
Db 67 TTGGATAGGTCATTTCTCTGCGAGTCTTGATCTCTCTTTAAATAGAGGA-AGTA 125
Qy 7429 AATCCACATGAGGAGTGAGGAGATCAGATCAACAGCTGATGATCATCTG 7488
Db 126 AATCCACATGAGGAGTGAGGAGATCAGATCATCAGCTGATGATCAACTG 185

```

```

Qy 7489 GTTTCGTGTTCCAGGGTCATCAGACTGAGGTTTCTGAGCATGATTCACCACTCCAGTC 7548
Db 186 GTTTCGTGTTCCAGGGTCATCAGACTGAGGTTTCTGAGCATGATTCACCACTCCAGTC 245
Qy 7549 TTGGGTCACAGAACTGACACCAATGACGAGAGAGACTCTTCTCTCAACAGACC 7608
Db 246 TTGGGTCACAGAACTGACACCAATGACGAGAGAGACTCTTCTCTCAACAGACC 305
Qy 7609 CTGAGCTTACAGGGGCTGACGTCGATGTTTCCCTTTCGCGCTGACAGAAAGCGGTT 7668
Db 306 TTGAGCTTACAGGGGCTGACGTCGATGTTTCCCTTTCGCGCTGACAGAAAGCGATT 365
Qy 7669 GTGCTCTGGCTCTGGGCTGCGGATGCGAGAACGCTGCTTCATCTACATCTTCAAC 7728
Db 366 GTGCTCTGGCTCTGGGCTGCGGATGCGAGAACGCTTCTCATCTACATCTTCAAC 425
Qy 7729 CTGGTCGCGGCGGACCTTCTCTTCTTCTAGCGGCGCAATTAATGTTGCGGTTACGCTC 7788
Db 426 TTGGCGGCGAGAGACTTCTCTTCTCTAGCGGCGGCGCTTAATTAATCTCTGTAGCTTC 485
Qy 7789 ATCAATATCCGCGCATCCCATCTCCAAATCTCTAGTCTGATGACCTTCCCTACTTT 7848
Db 486 ATCAATATCCGCGCATCCCATCTCCAAATCTCTAGTCTGATGATGTTTCTTACTTT 545
Qy 7849 ATAGGCTTAAGCATGCTGAGCGCCATGACACGAGCGCTGCTCATCTGTCGACC 7908
Db 546 GAGGCTGAGCTTCTGAGTGCCTGAGCACGAGCGCTGCTCATCTGTCGACC 605
Qy 7909 ATCTGTATCCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 7968
Db 606 ATCTGTATCCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 665
Qy 7969 GCCCTGCGCGCTGCGGAGATCTCTGAGAGATGTTCTGTGACTTCTGTTAGTGT 8028
Db 666 GCCCTGCGCGCTGCGGAGATCTCTGAGAGATGTTCTGTGACTTCTGTTAGTGT 725
Qy 8029 GCTGATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 8088
Db 726 GCTGATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 785
Qy 8089 TGTGTGTTCTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 8147
Db 786 TGTGTGTTCTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 845
Qy 8148 GAGA 8152
Db 846 GAAA 850

```

Search completed: December 2, 2004, 18:05:18
 Job time : 22726 secs

This Page Blank (uspio)